Hospital procedures concentration: how to combine quality and patient choice.

(Article begins on next page)

Doctoral Dissertation

Doctoral Program in Management, Production and Design (30^{th}cycle)

Hospital procedures concentration: how to combine quality and patient choice

A managerial use of the volume–outcome association

By

Elisabetta Listorti

******

Supervisor(s):

Prof. A. Alfieri

Politecnico di Torino

2018

Declaration

I hereby declare that, the contents and organization of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

Elisabetta Listorti
2018

* This dissertation is presented in partial fulfillment of the requirements for **Ph.D. degree** in the Graduate School of Politecnico di Torino (ScuDo).

Acknowledgements

The PhD journey has been quite a long journey, full of awaited and unexpected steps. I wish I could thank all the people that I have met during this journey, those who have walked with me for a long path as well as those who have just crossed my way.

The first thought goes to Arianna Alfieri, without whom I would have never started the PhD. Thank you for having believed in me, thank you for the many occasions in which you have given me your suggestions, and above all thank you for having always respected my preferences.

One of the first persons that has made me interested in research has been Andrea Matta, to whom I owe the core of the idea for my research. Thank you for having shared it with me.

Nothing would have been possible without the help of Paolo Sciattella, who has helped me in discovering the PNE, plus the secrets of the epidemiological work. Thanks also to all the PNE staff, for having hosted me during my PhD and having answered to all my curiosities.

The visiting in York has been fundamental for deepening the health economics techniques, but also for experiencing the life in a brilliant research centre. A strong thank goes to Andrew Street, who has guided me during those two months and on, teaching me how research can enrich policy analysis.

Even though these three years have often seen me puzzled, there is a group of people that has made living them easier, and often cheerful: thanks to Stefania, Danilo, Daniele, Gaetano, Matteo, Giovanni. A special thank to Erica, with whom I have shared the most.

Original thanks to in-laws, Daniela and Frank, for the support given during the production of the thesis.

The last thank goes all to my family, Matteo and Miriam. Thanks for having constantly reminded me that this PhD was neither the hardest nor the funniest challenge in my life.

Contents

List of Figures **viii**

List of Tables **xi**

1 Introduction **1**

2 Literature Review **5**

2.1 Literature on location and allocation problems 6

2.1.1 Allocation problems 6

2.1.2 Location and allocation problems 9

2.2 Literature on choice models 12

2.3 Literature on the volume–outcome association 14

3 Policy maker perspective **18**

3.1 Problem description 18

3.2 Model formulation 20

3.3 Health policy objectives 25

3.4 Case studies 29

3.4.1 Reallocation of interventions 30

3.4.2 Geographical distribution 32

3.4.3 Specialization thresholds 36

3.4.4 Hospital performance . . . 39
3.5 Limitations . . . 44

4 Patient perspective 47
4.1 Problem description . . . 47
4.2 Model formulation . . . 49
4.3 Case study . . . 53
4.3.1 Data . . . 58
4.3.2 Descriptive Statistics . . . 61
4.3.3 Results . . . 67
4.4 Limitations . . . 72

5 Integrated perspective 73
5.1 Problem formulation . . . 75
5.1.1 Policy maker model . . . 75
5.1.2 Patient model . . . 79
5.1.3 Quasi-integrated approach . . . 83
5.1.4 Fully-integrated approach . . . 89
5.1.5 Comparative analysis . . . 94
5.2 Case study . . . 98
5.2.1 Policy maker solution . . . 98
5.2.2 Patient solution . . . 100
5.2.3 Quasi-integrated solution . . . 103
5.2.4 Fully-integrated solution . . . 105
5.2.5 Comparative analysis . . . 106
5.3 Policy Implications . . . 110

6 Conclusions 114

References 118

List of Figures

3.1 Hip fracture; association between 30-day mortality and volume of activity by facility, Italy 2011 [1] 26

3.2 Function M_1 behavior in hip fracture surgery, with total demand equal to 300 27

3.3 Function M_1 behavior in hip fracture surgery; total demand equal to 300 (a), multiple total demands (b) 28

3.4 M_2 behavior in hip fracture surgery (total demand equal to 300); PNE mortality curve case (a) and monotonic linear curve case (b) . 28

3.5 M_3 (a) and M_4 (b) objective function behavior in hip fracture surgery (total demand equal to 300) 29

3.6 Volume allocation between hospitals in different configurations . . . 31

3.7 Comparison between average mortality (M_2) and total mortality (M_1) of each type of intervention in the three configurations 32

3.8 Case study 1. Geographical distribution of volumes for bladder cancer interventions performed in Piedmont in 2013 [2] 33

3.9 (a) (b) Case study 1. Distribution of volumes for bladder cancer interventions performed in Piedmont provinces in 2013 [2] 34

3.10 Case study 1. Bladder cancer, analysis of the association between 30-day mortality and volume of activity by facility, Italy 2011 [1] . . 35

3.11 Case study 1. Total mortality and average mortality for different minimum hospitals imposed per province 36

3.12 Prostate cancer and colon cancer; analysis of the association between 30-day mortality and volume of activity by facility, Italy 2011[1] . . 37

3.13 Case study 2. Comparison among optimal allocation of volumes (threshold 30) and volumes actually performed in 2013, Alessandria province 39

3.14 Case study 2. Comparison among objective functions' values with thresholds 0-40 40

3.15 Effective risk-adjusted outcomes (2013) compared to PNE mortality curves, for lung (a) and colon (b) cancer interventions 41

3.16 Interpolation of the average absolute value deviations, for lung (a) and colon (b) cancer interventions 42

3.17 PNE mortality curves with deviation bounds, for lung (a) and colon (b) cancer interventions 42

3.18 Case study 3. Actual outcomes compared to mortality curves, and performance coefficients for three hospitals of the province of Turin, for lung (a) and colon (b) cancer interventions 44

3.19 Case study 3. Actual volumes of colon and lung cancer interventions performed in three hospitals in the province of Turin in 2012 [3] (a) and optimal allocation (b) 45

3.20 Case study 3. Comparison among outcomes in actual and optimal configuration, for three hospitals of the province of Turin 46

4.1 Colon surgeries performed by hospital, 2014 54

4.2 Hospitals location in 2004 (red and yellow dots) and 2014 (red dots) 55

4.3 Patients' distribution in Piedmont 55

4.4 Number of wards operating colon surgery 56

4.5 Location of hub hospitals 56

4.6 Volume of activity in hub hospitals in Piedmont 57

4.7 Clinical pathway for colon cancer patients in Piedmont 58

4.8 Hospitals' characteristics over time 62

4.9 Hospitals' volumes over time 62
4.10 Patients' characteristics over time 64
4.11 Percentage of bypassers from 2004 to 2014 65
4.12 Percentage of patients choosing the nth closest provider, 2014 . . . 65
4.13 Differences among performance of hospitals chosen by bypassers/no bypassers 67
4.14 Additional distance travelled for different categories of patients . . . 69
4.15 Results on year interaction terms 70

5.1 Flow chart describing decision maker and patient models 73
5.2 Solution algorithm for the quasi-integrated approach 87
5.3 Description of the algorithm of the fully-integrated approach 91
5.4 Number of colon cancer interventions performed in hospitals in Piedmont, 2014 98
5.5 Comparison among volume distributions of actual and strategical solutions 100
5.6 Comparison among outcomes of actual and strategical solutions . . 101
5.7 Patients who are assigned to their nth nearest hospital, quasi-integrated approach 103
5.8 Additional distance travelled, quasi-integrated approach 104
5.9 Mortality for operational solutions 106
5.10 Comparison among outcomes of actual, strategical, operational and fully-integrated solutions 108

List of Tables

3.1 Notation used for the mathematical formulation of the decision maker model 21

3.2 Example of configurations to be evaluated with the average range objective function (M_4) 29

3.3 Case study 2. Volumes of colon and prostate cancer interventions performed in Alessandria province's hospitals in 2013 [2] 38

4.1 Descriptive statistics about hospitals' characteristics 61

4.2 Descriptive statistics about patients' characteristics 63

4.3 Differences among possible patient choices and actual patient choice 66

4.4 Results from the conditional logit 68

5.1 Notation used for the mathematical formulation of the policy maker model 77

5.2 Notation used for the mathematical formulation of the patient model 81

5.3 Notation used for the quasi-integrated approach 84

5.4 Key performance indicators for the comparative analysis 107

Chapter 1

Introduction

The main focus of this PhD thesis is the use of managerial tools in the healthcare sector. In particular, the principal topic we focus on is the volume-outcome association, a relationship that has been empirically identified in medical specialties. According to this relationship, there exists a positive association between the number of interventions (the so called volume of activity) performed by a facility and the quality of clinical outcomes, measured in terms of patients' health conditions.

The volume-outcome association has been identified back in 1979 [4], and it has been particularly documented in the last two decades for a variety of interventions and different outcome measures [5, 6]. All the studies mainly reveal that there is a positive effect of volume on outcomes for each medical procedure, although its extent varies depending on the clinical area itself. The observed trend can be explained by two main factors: (i) on a hospital level, the structure by which care is organised is likely to be poorer in low volume hospitals, which might lack consistent processes for postoperative care or for dealing with postoperative complications [7]; (ii) on a personnel level, outcomes may also be related to the familiarity of the staff with the treatment [8, 9]. Despite the number of studies focusing on it, the volume–outcome association still raises interest, due to the persistence of low volumes performed in healthcare facilities, in particular in Italian hospitals.

Our starting point is the National Outcome Evaluation Program (PNE), a project sponsored by the Italian government that each year, from 2012, reports hospitals clinical performances with the objective to assess healthcare service quality levels [10]. Furthermore, in 2012 and 2017, it published two works showing the occurrence of the volume-outcome association for several procedures performed in Italian hospitals [1, 11].

While many researchers have focused on the existence of the volume-outcome association from a clinical perspective, this PhD project deepens the volume-outcome association from a managerial perspective, by including it in a planning problem. The planning problem analysed consists in the decision of how to distribute volumes of activity among wards of hospitals operating in a same geographical area. In particular, among the different specialties, we consider surgery wards, since better results for higher volumes are especially plausible for this case. In fact, high-volume hospitals may have more surgeons who specialize in specific procedures (thus gaining learning from experience [12]) and more consistent processes for all the phases of care [7]. Our ultimate objective is to exploit the information contained within the volume–outcome association and, as a consequence of the existing link among volume and outcomes, to reach an optimal planning for hospital wards. In this way, the reorganization of hospitals operating in a territory (planning decision) translates into the improvement of healthcare organization outcomes (clinical result). In the following, we will refer to the volume-outcome association also with the term *mortality curve*, since it can be translated into a function linking the mortality and the number of treated patients.

We take as reference healthcare system the Italian National Healthcare System (Servizio Sanitario Nazionale, SSN), a public health system that provides universal coverage for comprehensive and essential health services [13]. The SSN is based on three main principles: universality, equality and equity. By universality, we mean that health services are provided to the whole population, so that the SSN encourages all the citizens to care about their physical and psichological health status. By equality, we mean that no difference among citizens is made depending on their individual, social or economical characteristics. By equity, we mean that people with the same health needs are guaranteed the same access to the healthcare services.

The formulation of our problem varies depending on which actor is considered. In the SSN, there is a central decision maker, the commissioner, in charge of guaranteeing the compliance with the above mentioned principles. It is represented by an institutional figure at the national level, i.e., the Ministry of Health. However, all the administrative levels (e.g., Regions, municipalities, etc.) have to collaborate in order to guarantee health quality to all the citizens. Hence, we can think of a commissioner at each layer of the system, which is responsible for the population health.

Beyond the commissioner, other actors compose the Italian healthcare system. In particular, three other categories are involved in planning problems: providers, physicians and patients. Commissioners demand to providers to supply healthcare services. Providers (i.e., hospital administrators) answer through the supply of the requested services. Medical staff (surgeons, anesthetists, nurses, etc.) are the experts who deal with patients, who in turn receive the service. It should be noticed that there is no constraint enforcing patients to choose a specific hospital where to be treated, and no patient is forced to receive healthcare services. Each actor has its own interests and perspectives [14], and therefore it is relevant to keep into consideration their different behaviors and interactions.

Since the allocation of operation volumes to healthcare structures is a strategic decision that deals with territorial healthcare configuration and people health needs, we initially take the perspective of the commissioner, who is the first actor involved in this decision process. All the other actors will face the consequences of such strategic choice: providers will have to adapt the capacity of their structures to the new planned demand; medical staff will have to arrange new shifts and work organisation; patients will face new openings/closures of hospitals and will have to choose where to be treated. Among them, we reckoned as particularly worthy of attention the patients' perspective, since their behaviour can alter the whole commissioner plan.

The thesis is structured as follows. Chapter 2 summarizes the relevant literature. The chapter is organized in two sections dedicated to the two main fields of studies we refer to, namely location and allocation problems (from the health management literature) and choice models (from the health economics literature). Moreover, a section of the chapter reports the state of the art of the researches that have been conducted on the volume–outcome association.

Chapter 3 is dedicated to the policy maker's perspective. We take the point of view of the commissioners, i.e., that of planning the volume to be allocated to each hospital, and we propose an approach (based on mathematical programming) to determine the number of interventions to be strategically allocated to surgery wards, given several constraints related to hospital capacity, demand satisfaction and epidemiological concerns. Concentration vs. scattering of interventions among healthcare structures are explored in terms of quality and equity offered to the whole population. The proposed approach is tested on four case studies taking into account real life factors (such as reallocation of interventions, geographical distribution of hospitals, volume threshold constraints, and dissimilarities among hospital performances), and results are compared with real data from the PNE.

Chapter 4 focuses on patients' perspective. Specifically, we analysed patients' choice, in terms of hospital where they have decided to be treated, together with the list of hospitals that were available to them. By using the econometric methodology of the conditional logit, we modeled the trade-off faced by patients between hospitals' characteristics, i.e., distance and quality. Eventually, we applied the choice model to Hospital Discharge Data for colon cancer patients in Piedmont from 2004 to 2014, showing patients' revealed preferences. Results shed some light on how patients can react to facility specialization or closure, depending on demographic, social and clinical factors.

Chapter 5 gathers the two perspectives and merge them. The objective is to support planning decisions that (i) are effective in terms of better health outcomes and (ii) guarantee patients' choices to respect the volumes that have been strategically planned. To this aim, we explored two distinct approaches. The first approach enriches the one proposed in Chapter 3 with the commissioner point of view, by adding constraints involving patients, e.g., the maximum distance they are willing to travel. The second approach, instead, aims to fully integrate patients' and policy maker's perspectives, by inserting predictions on patients' behaviour within the decisional process of the policy maker. Eventually, results from all the approaches are compared, in terms of organizational quality and population health.

Chapter 2

Literature Review

Planning problems in healthcare systems have received great attention in the last decade. European policy makers' agenda include the need for a territorial reorganization of healthcare services provision, and, at the same time, patients advocate for quality and certainty of clinical outcomes. Changes in the composition of the population, in the medical techniques and in the managerial frameworks are only some of the aspects that increase both interest and difficulty in the topic. Hence, there exists a variety of studies on planning problems, which analyse different approaches and perspectives.

As stated in the introduction (Chapter 1), in this thesis we study the reorganization of hospitals operating in a territory (in terms of volume allocation), with the objective of improving their clinical outcomes. We examine this topic from two perspectives, namely (i) the policy maker and (ii) the patient point of view.

The problem, as seen from the policy maker perspective, is a planning problem. As such, it includes decisions about the number of hospital services to be offered, their location, and details in terms of the volumes of specific allocated resources [15]. The most used approaches to such kind of problems can be clustered into two main areas, which can be labelled as *location* and *allocation*. The main contribution to these areas, in the health management literature, are discussed in Section 2.1.

Instead, the problem, as seen from the patient perspective, is a choice problem. In fact, patients have to decide in which hospital, out of the hospitals that are available on the territory, they want to be treated. Section 2.2 reports the main contributions on choice models, in the field of health economics.

Eventually, Section 2.3 deepens the arguments that have been debated on the volume–outcome association.

2.1 Literature on location and allocation problems

The importance and significance of planning decisions in healthcare have been abundantly recognized in the health management literature of the last decade [16]. In fact, managerial tools and planning techniques can represent an answer to the growing need of homogeneous levels of clinical quality and to the existing pressure among hospitals due to limited budget and resources. The aim of regional healthcare planning has been defined as that of providing figures about the number of hospital services to be offered in the future, their location, and details in terms of the volumes of specific allocated resources [15]. This aim can be approached referring to the two main areas of *location* and *allocation*, as poor location and/or allocation decisions can affect the health of the population in a territory [17–19]. Considering the healthcare management field, location problems have been considered always together with allocation problems, while allocation has been considered also by its own.

2.1.1 Allocation problems

The allocation problem can be generally defined as to how optimally assign a limited amount of resources to a set of demand points, in order to achieve a given objective. It is usually considered together with capacity planning, since the two issues are related to each other. Papers in this field mainly differ for i) the resource to be allocated, ii) the level of details of the study, and iii) the objective function.

Considering the resource, the majority of the studies are related to the decision about the number of beds (e.g., [20–25].). Others considered the number of appointments to be scheduled [26], or the time available per day of the operating room [27], and only few focused on the number of patients that are treated [28]. Some studies also considered the simultaneous allocation of different resources. For instance, in [21], the flow of patients is simulated through user-defined care-units, which may represent a ward, a specialty bed-pool or the hospital as a whole. In [29], authors included as resources beds, operating theatre capacity and nursing staff. In [30], patients and operating room time are taken into account. Both operating rooms and working hours have been accounted for in [31].

Instead, about the level of detail, it mainly depends on the time horizon that is used. The strategic level involves a long term horizon, through planning the monthly or yearly capacity (e.g., [27, 20, 21, 23, 22, 24]. On the opposite side, analysis conducted on the operational level concentrate on the capacity that is needed/allocated daily [21, 31]. Tactical level concerns something in between the two, as the generation of a master surgical schedule ([29, 32–35]) and the definition of the mix of patients to be admitted. Interestingly, some studies have taken into account multiple levels. For instance, in [36], authors developed a three stage process that involves all the three levels: the allocation of operating room time to surgical specialties at strategic level, the development of a master surgery schedule at tactical level, the scheduling of individual patients at operational level. Usually, studies that focus on working hours are operational, while those on beds are strategic.

Finally, the most used objective functions in healthcare management are profit, time, utilization and quality. Some researchers included multiple objectives in their study, such as [31], with the trade-off among three performance criteria: wait before being on schedule, scheduled procedure start-time reliability, hospital profits.

As an example with profit, the work [36] considers, within the case mix planning problem, the maximization of the overall financial contribution.

Examples for objectives on time can be found in [33] and [26], who aimed to offer to patients a timely appointment, which guarantee short appointment dates and avoid long backlogs. Similarly, in [28], authors considered completion times in Accident and Emergency departments. In [30], instead, authors deepened the positive effects achievable in wait thanks to both increased capacity and to the effective use of current funds.

About utilization, two key hospital indicators were used in [21]: bed occupancy (i.e., the proportion of time that a bed is occupied) and refusal rate (the proportion of referrals that is refused because no bed is available for an arriving patient). In [22], the objective was the minimization of the number of patients transported outside the region due to overloaded capacities. In [29], authors aimed to better performance on target utilization levels of resources, by minimizing the weighted sum of expected under and over utilization. Service level, occupancy level and admission rate were used in the objective function in [23], while occupancy level and refused admission rate are used as key indicators in [21]. Other works focus on combinations of the mentioned objectives [20, 36, 32, 24, 25, 27].

It has to be mentioned that, apart from profits, both time and utilization are related to clinical quality. For example, frequent bed shortage (that we consider under the utilization area) can be considered as a proxy of low quality. In fact, solutions adopted in the case of bed shortage are either transferring a patient to another hospital/region, or postponing a planned operation, or releasing another patient earlier – all solutions that have serious drawbacks in terms of patients' health [22]. However, time and utilization act as indirect indicators of quality, since they do not quantitatively measure it; rather, they associate quality with some measured managerial inefficiency.

Quality of the care received by patients, instead, can be considered explicitly through the measurement of patients' health, even if only few studies included it as objective. As an example, in [31], they assumed quality to be constant. Nonetheless, service quality aspects in healthcare delivery play an important role in patient satisfaction [29].

In this thesis, we considered the allocation of the number of interventions, required by patients that live in a territory, among the hospitals of the same territory. This problem is a strategic problem and the criterion of the allocation is based on the volume—outcome association, since the ultimate objective, which originates from using the perspective of the policy maker, is to increase the quality of the outcomes. The most relevant element that distinguishes our work from the previous literature is the use of a direct and explicit measure of quality. In fact, our priority is to maximize the health conditions of the population, which we directly monitor through the measure of mortality for surgical interventions.

The main criticality of the allocation problems is that, once the resource is allocated to the healthcare structure, it is assumed that it will be used efficiently. This is realistic in the case of operational problems, where the efficiency of the process entirely depends on the medical staff (for instance, the organization of the shifts among colleagues). However, on a strategic level, the use of the resources depends also on patients' behaviour. An example is the allocation of patients to facilities, which is not usually controlled by the policy maker [37], rather by the preference of people who choose to seek their service. This is what happens in the SSN, where the ultimate choice on where to be treated rests with the patient. In this case, patients can transform the allocation solutions formulated by the policy maker with their hospital choices. In fact, even if a hospital is planned to perform a high number of interventions, only a high number of patients' positive choices for that hospital will confirm the planned volume.

2.1.2 Location and allocation problems

Starting from a set of demand locations and a set of candidate facility locations, the function of location/allocation problems is to identify (i) the geographical area where the structures should be placed, and (ii) how the demand should be divided among the structures. The identification process is guided by a specific objective function, which is related to the way the demand is fulfilled. These problems focus only on a strategic level, since they always involve the total reconfiguration of services.

In [17], several location models are reviewed, while [38] provides a comprehensive overview for methods and criteria, together with a summary of the main challenges and solutions of these problems. A review of location and allocation models restricted to developing countries can be found in [39].

Usually, the location-allocation models differ among each other along two main dimensions. The first one is the type of structures they refer to. Studies examined multihospital systems [40, 41], preventive healthcare services [37, 42], long-term care services [43], emergency medical service systems [44], maternal and perinatal assistance [45], transplant systems [46], blood banks [47], or general healthcare facilities [48].

The second dimension that distinguishes location and allocation models is the objective that drives the process of location. The three main objectives used refer to three classic formulations of facility location allocation problems: the P-median problem, the covering problem and the center problem [17].

In the P-median problem, the objective function minimizes the total travelling distance from all clients to their closest serving facilities, by minizing either the average distance (average performance) or the sum of distances (overall performance) [42]. The idea is that when the average (total) distance decreases, the accessibility and effectiveness of the facilities increases [44].

The covering problem, instead, focuses on the coverage, i.e., the ability to serve more people within a target area. A demand at a node is considered as covered by a facility located at some other node if the distance between the two nodes is less than or equal to some exogenously specified coverage distance [17]. Covering problems can have two objectives: either to minimize the number of facilities while providing coverage to all the demand nodes, or to maximize the coverage provided the number of facilities is pre-specified.

The center problem, for a given number of facilities, identifies a location arrangement that minimizes the maximum distance while requiring coverage of all clients [42]. By attempting to minimize the worst performance of the system, it addresses situations in which service inequity is more important than average system performance [44].

Applications of these three classical problems can be found in many recent papers, like [44, 48]. Besides these standard problem structures, researchers have experimented either enriched versions of three objective functions [49–51], or combinations of them, or with tasks that differ from distance and coverage. Very often, studies address multi-objective problems that cover the multiple priorities of the stakeholder involved [52]: surely coverage [42] and distance [45], but also efficiency, accessibility [42], load imbalance [45], time [46] and cost [41, 47]; fairness and equity [46, 47]. In [53], the use of simulation tools allows the user to monitor the consequences of the geographical distribution and organization of the services in terms of multiple outputs.

An alternative to multi-objective problems is to transform objectives into constraints. Constraints are normally resource limitations, and are related to demand, efficiency, capacity, distance and cost [38]. For instance, in [37], an approach based on the covering problem is proposed, where each facility needs to have a minimal number of clients; in [40], authors minimize total cost while also enforcing a target level of patient service; in [43], they considered the objective of minimizing the maximum load of facilities under the constraints that demands for the care are assigned to the closest facilities.

To the best of our knowledge, the only paper that considered as a benefit the maximization rather than the minimization of the workload, is the one from [37]. In fact, the authors stated that the workload of a facility needs to be greater than a certain threshold so as to ensure the quality of services. However, due to this consideration the volume becomes part of a constraint, rather than included in the objective function.

It should be noted that location and allocation problems do not include the criteria of the linkage among demand and supply points. In other words, after having set the supply locations, the solutions of these problems do not explicitly assign specific demand points to specific supply points. Some of the location and allocation problems implicitly assumed that people would seek services from the closest facility [37]. Others assumed that patients would seek the services from the facility that satisfy some criteria, such as minimum expected total time [24]. Others did not specify any criteria into the solution algorithms for the decision of the assignment [42].

In this thesis, a problem that might resemble location and allocation problems emerges when we considered policy maker and patient perspective at the same time. However, many differences prevented us from addressing it as in the previously cited papers. First of all, the location part does not appear to be from greenfield, meaning that we never assume to have a plain space that needs to be filled with hospitals. Instead, our problem starts from the brownfield configuration of healthcare facilities that are already placed and open on the territory. No facility has to be newly established, rather our decision is in terms of which facility may be closed.

Secondly, as explained in Section 2.1.1, our approach in the allocation process differentiates because of the objective. The ultimate objective considered in this thesis is the quality of the clinical outcome. We define quality in terms of patients' mortality. To the best of our knowledge, there is no paper in the management literature that address the allocation problem from this perspective. Moreover, our approach is innovative in considering the outcome through the use of a managerial factor, i.e., the volume of activity. Even though the workload has been considered in the literature, no work has exploited it for its link with the clinical outcome. This innovative approach is allowed by the use of the volume-outcome association, which has received wide attention in the medical literature. Section 2.3 reports some of the most recent review and studies on the topic.

Third, the consideration of patients' welfare. Instead of including in the objective a weighted or averaged measure of distance or quality perceived or satisfaction, we insert the outcomes of patients' decisional processes by considering their individual preferences. By doing so, we relax the assumption of patients seeking services always from their closest facility. The idea is to give importance to individual patients, not only as parts of the planning process, but as people willing to be treated with the highest service quality. Since the literature on health management does not make use of this approach, we explored the health economics knowledge. Indeed, many researchers in this field have deepened patients' perspective in the planning process, whose main part is the choice of the hospital to be treated in. Hence, section 2.2 is dedicated to the health economics literature about patient choice.

2.2 Literature on choice models

Within the health economics literature, models representing patient choice have received a great consideration. Indeed, this reflects the interest paid by all the healthcare stakeholders on the topic. In particular, governments' policies are increasingly fostering patient choice among providers, therefore favouring their mobility [54]. The final aim of these policies is to boost improvements on the offer side (i.e., hospitals), by acting on the demand side (i.e., patients). In fact, if patient choice was firmly driven by quality, high quality hospitals would be awarded by the increase of healthcare services required, whereas lower quality hospitals would be forced to improve in order to stay in business or financial balance [55].

In general, the papers in this field take the territorial context into account: UK [56–59], US [60–63], Italy [64], Netherlands [65–67], Japan [68]. The country patients come from is of interest because of the different campaigns used by governments in order to encourage patients to choose. For example, in UK, starting from 2008, patients have the possibility to choose to be treated in any of the national providers. At the same time, the government has increased communication channels to spread information on hospitals' performance, e.g., a website has been introduced to publish easily accessible data. As far as Italy is concerned, patients have always had free choice on the provider, even though it has never been entirely supported by a wealth of information enabling patients to decide wittingly. The only government sponsored website is originated by the project called National Outcome Evaluation Program (PNE), whose aim is to measure the outcome variability among providers. However, accessibility to the website has not been documented yet.

Beyond the geographical location, papers focus on the different elements that affect patient decisional process. Researches have shown that a complex interplay between hospital and patient characteristics determines patient choice [69]. As for the hospital level, all the studies considered two factors: distance and quality. In fact, distance between its location and people residence has been shown to discourage patient choice. This happens because distance is positively related to the travel time, which has as a consequence a greater effort required to reach the hospital, both for the patients and their family and friends who want to visit them.

At the same time, higher outcome quality levels strongly motivate patients to bypass their nearest provider in order to choose a better one. The quality of care is valued because it may lead to better health outcomes, and because it may improve the welfare of the patient during the process of treatment indeed [60]. Different measures of quality can be used, either based on production inputs, or on patient outcomes [60]. As for the input, studies have focused on range of specialized services offered (e.g., the presence of ICU, the emergency department, the number of specific complex procedures performed) [68], hospital size [60], number of beds [57], teaching status (i.e., if the hospital provides medical education and training), number of hospital admissions [70], volume of interventions [62, 71], physician characteristics [72].

As for the output, studies have used waiting times [56, 73–75, 65, 76], readmission rates [66, 56], length of stay [68], mortality [57, 75], complication rates [60], number of patients transferred out of the hospital [61], official ratings [66, 77], patient-reported outcome measures [58]. Some papers have concentrated on more pathology-specific outcome measures [57, 78, 79].

As for the individual level, studies differ depending on (i) demographic characteristics, like age and sex [59]; (ii) social and economic conditions, like frailness, urban/rural [56], social interactions [64], income [62], employment status [78], and health-deprivation measures [57]; (iii) clinical conditions, such as medical pathology [65, 63], the severity of the illness [73], and the number of comorbidities [58]. Patients' characteristics can affect the way they trade off quality and distance. For instance, since age can be viewed as a proxy for the frailness or the ability of the patient to travel, younger patients can be more willing to travel than older patients.

Other sources of difference among studies origin from the clinical area: hip replacement [56–58], angioplasty, coronary artery bypass graft (CABG) [73, 75], heart attack [60], colon cancer surgery [68], breast cancer surgery [71], obstetrics care [62], angioplasty [66], kidney transplatation [78] and family doctor [59]. This difference raises interest because (i) depending on the specific treatment, patients can be elective or emergency; (ii) different treatments involve different clinical pathways, which affect patient decisional process. Patients that urgently seek medical care are supposed to consider hospital quality as a more important determinant of choice [60]. On the other side, patients that need a planned procedure have time to realistically make a choice, and their choice is more likely to be informed, for example, through conversations with peers [57]. At the same time, for all the urgent patients, the time spent between the occurrence of the clinical event (e.g., a heart attack) and the delivery of treatment affects the efficacy of the treatment [60]; hence, distance becomes crucial for their choice.

Eventually, but most importantly for the purpose of our work, differences emerge on the rationale of the study, which deeply affects the use of results that is implemented or suggested. In the literature, choice models have been mainly used to examine changes in the behaviour of patients in response to real or simulated government policies [54]. These policies can refer to territorial changes, such as hospital mergers [57], hospital closure [60, 62], increase in the number of rivals [56]; or to structural changes, like the increase in personnel or technical tools [60]; or to policy changes, like removal of constraints on patient choice [73]; or simply to actual configurations [58, 68, 70].

In our work, as in the mentioned literature, we aim at quantifying patients' preferences based on their past choice. However, we also included how patients choose in the planning problem of the policy maker, when the two actors' perspective are considered in an integrated view.

2.3 Literature on the volume–outcome association

As stated in the introduction, the principal topic we focus on is the volume–outcome association, a relationship that has been empirically identified in medical specialties. According to this relationship, there exists a positive association between the number of interventions performed by a facility and the quality of clinical outcomes, which in turn means better patients' health conditions. The volume–outcome association has been identified back in 1979 [4], and it has been particularly documented in the last decades for a variety of interventions and different outcome measures. The interest is raised by the relevant impact on effectiveness of healthcare represented by hospital or physician volume [1].

Studies on the volume–outcome association include papers that consider single procedures [6], papers that refer to groups of interventions [7], meta-analysis [5], systematic reviews [8, 80], and systematic review of systematic reviews [1, 11, 81].

As far as Italy is concerned, the National Outcome Evaluation Program research group in 2012 published a document aiming to (i) identify clinical conditions or interventions for which an association between volume and outcome had been investigated or proved in the literature, (ii) analyse Italian health providers in terms of the volume of activity and of the association between volume of care and outcomes [1]. From the systematic review of systematic reviews, evidence of a positive association between volumes and intrahospital/30-day mortality was demonstrated for 25 clinical areas. Among these conditions, for 14 procedures (where national data had sufficient statistical power) authors observed a positive association between volume and outcome of care from data on Italian health providers. The study has been updated at the end of 2017 [11]. The new systematic review of systematic reviews has confirmed the volume–outcome association for 34 clinical areas. The analysis on Italian data has been conducted for 16 procedures, where the association between volume and the outcome of care has been observed.

All these international studies mainly reveal that there exists a positive effect of volumes on outcomes for each medical procedure, although its extent varies depending on the clinical area itself. In other words, the shape of the found relationship differs among different interventions, with heterogeneous slope of the curves. For example, less frequent clinical interventions, which usually are the most difficult, show largely better outcomes if volumes performed increase. In general, analyses show a strong improvement in outcomes in the first part of the curve (from very low volumes to higher volumes) for the majority of the studied conditions [1].

The explanation for the observed trend originates from two main factors [82]. On a hospital level, the structure by which care is organised is likely to be poorer in low volume hospitals, which might lack consistent processes and greater resources for postoperative care or for dealing with postoperative complications [7]. Moreover, high-volume hospitals are thought to benefit from a multidisciplinary team work approach, local availability of other specialist services and more active involvement with research [83].

On a personnel level, outcomes may also be related to the familiarity of the staff with the treatment, leading to the *learning by doing* mechanism [84]. In fact, the greater experience of the surgeons that perform a high volume of interventions should lead to improvements in the pre- and intraoperative decision-making process, case selection and surgical technique [5].

At the same time, there could be reverse causality in the volume—outcome relationship: physicians and hospitals with better outcomes may attract higher volumes of patients (the so called *selective referral*) [85], even though some studies have found that the direction of causality is primarily from volume to outcome [86]. All in all, more efforts should be devoted to investigate not only the presence of the volume–outcome association, but also the underlying mechanism [87].

Despite the number of studies focusing on it, the volume–outcome association still raises interest, since there is a great number of hospitals, especially in Italy, which continue on performing low volumes of activity.

In this thesis, differently from the above cited documents that focus on the existence of the volume—outcome association, we study an innovative use of it, which allows to merge the managerial and clinical perspectives. Our ultimate objective is to exploit the link among volume and outcomes that is provided by the mortality curves, in order to suggest an optimal planning that results in better clinical results. In this way, the reorganization of hospitals operating in a territory (planning decision) translates into the improvement of healthcare organizations outcomes (clinical result).

Actually, different applications of the volume–outcome association have already been suggested. For example, in US, the Leapfrog group (i.e., a consortium of healthcare purchasers) recommended minimum hospital volumes for some interventions [88]. However, for many procedures, the improvement in outcomes remains gradual or constant with the increasing volume of care, thus making harder the identification of threshold values beyond which the outcome does not improve further [1]. Moreover, the lack of information generally valid prevents from any precise numerical recommendation [81]. This is why the set of a threshold remains controversial. Another action that has been suggested is to educate patients, by providing and explaining to them data on the hospital volumes and thus promoting selective referral [89]. However, great effort should be spent in guaranteeing that measures of performance quality are fully understood by patients.

Our approach appear to be promising since it analytically explores other possible uses of the volume–outcome association. We include the mortality curve in a managerial framework that, by accounting for multiple factors that are relevant to the involved stakeholders, suggests a tailored optimal solution.

Chapter 3

Policy maker perspective

This chapter considers the perspective of the decision maker in dealing with a planning problem, i.e., the optimal allocation of surgical interventions to several hospital structures, based on the objective of clinical quality. The main contribution of this chapter is the exploitation of the volume-outcome association within the decisional process of the decision maker. In fact, the allocation of different volumes of surgical interventions to the hospitals of a given region leads to different territorial configurations of the healthcare services. By using the volume-outcome association, it is possible to evaluate clinical outcomes of each territorial configuration. Since decision makers are mainly interested in guaranteeing health quality, the proposed tool is helpful to compare dimensioning decisions and opt for the best one in terms of patients' health.

The content of the section is published in [90].

3.1 Problem description

The decisional problem of the policy maker considers a geographic region with a given number of hospitals, which perform several surgical interventions. The final aim is to allocate the number of interventions required by patients to the set of regional hospitals, in order to improve healthcare quality.

The allocation problem can be formalized by using a mathematical programming model in which the objective function represents the main objective of the policy makers related to the population health conditions, and constraints define the limitations caused by structural and epidemiological reasons. The decision variables are represented by the volume of activity of each intervention that is to be assigned to each hospital.

Thanks to epidemiological studies, the total number of patients who require to be operated in a given year, for each surgical speciality, is assumed to be known in advance. Constraints are related to (i) the total cost of the performed interventions, which has to be lower than the regional funding; (ii) the risk of under/over-treatment (i.e., not operating patients who need it, or to operate patients not needing it, respectively); (iii) the capacity of each hospital (in terms of performable interventions).

The objective function, which has to drive the solution (i.e., the volume allocation) towards the best configuration, refers to the mortality risk, defined as the ratio between the number of patients who died during a given time interval and the total number of patients treated during the same time interval. Mortality risk is the key input parameter of our problem and can be considered as the patients' probability to die after the intervention. It strictly depends on the type of surgery; nonetheless, mainly four groups of factors can further affect this probability: 1) patients' characteristics, such as age, sex, comorbidities; 2) surgeon, medical staff and facility experience (i.e., the positive impact of volume on outcome); 3) surgeons' talent; 4) uncertainty, in terms of the variety of accidental and uncontrolled elements, which prevails in small hospitals. Making explicit the dependence of mortality from the intervention type and considering risk-adjusted mortality (that absorbs the first factor impact, i.e., the patients' characteristics), the mortality risk can be considered mainly a function of the number of treated patients (the mentioned second factor). In fact, all the surgeons should be able to perform successfully, and interventions and accidental elements should not, in general, prevail on the other factors.

Even though the policy maker objective is apparently unique, i.e., the improvement of patients' health conditions, this goal can actually be translated into different health policy objectives, each of them enlightening a specific aspect. In fact, different indicators can be used to evaluate health policy results [91], but none of them is better than the others, since each of them can enlighten different (if not opposite) views on health needs [92]. In this chapter, four objective functions are considered. Namely, the total mortality, the average mortality, the mortality variance, and the average mortality range have been examined. Each of these objectives is expressed in terms of the patients' mortality and highlights different decision maker's interests, which are centred on the main targets of equity and quality.

Section 3.2 is dedicated to the model formulation. From that, mainly two groups of analysis were performed. First of all, different mathematical formulations of the objective function were tested. Configurations resulting from the use of each objective function were compared, in order to understand the trade-offs among them, as shown in Section 3.3. The model was applied to four case studies considering the reconfiguration of Piedmont Region surgery wards in Section 3.4. For each case study, the optimal solution was compared with the actual configuration (data from PNE).

3.2 Model formulation

In this section, we propose a mathematical formulation of the problem. To better understand the impact of the volume allocation on the objectives of equity and quality, we consider a deterministic setting. In the reality, the planning problem has a stochastic nature. Dimensioning decisions are in fact based on forecasts, and even when they reproduce real population needs, patients still have the possibility either to choose in which hospital they want to have surgery or to choose not to have surgery. This possibility causes a difference between the forecasted number of operations (what is strategically planned) and the performed ones (what actually occurs) for each hospital. However, in this formulation, we deal with the deterministic case, i.e., we assume that what we plan will surely occur.

Let J be the total number of hospitals that are in the considered geographic area ($j = 1, \cdots, J$) and S the types of surgical operations ($s = 1, \cdots, S$). The decision variables of the problems are the x_{js} variables, which represent the volume of surgical operation of type s allocated and performed in hospital j.

The actual request of operations for surgery s in the given geographical area is denoted by d_s. The *mortality risk* of patients receiving operation s is indicated by $m_s(x_{js})$, to make explicit the dependence of the mortality on the volume of performed operations. These elements are input parameters, as also the marginal cost, c_s, for each surgery type s, the fixed regional funding b available to pay for all the performed operations, the required capacity for each operation $cap.op_s$, the

Table 3.1 Notation used for the mathematical formulation of the decision maker model

Indexes	
Name	**Definition**
j	Hospital
s	Type of surgical operation
Decision variables	
Name	**Definition**
x_{j_s}	integer variables, which indicate the total volume of the surgical operation of type s performed by hospital j
Parameters	
Name	**Definition**
J	Total number of hospitals
S	Total number of types of surgical operations
c_s	Marginal cost for surgery of type s
b	Available regional funding
$m_s(x_{j_s})$	Mortality risk of patients receiving operation of type s in hospital j
d_s	Number of operations of type s required in the geographical area
$cap.op_s$	Required capacity for each operation of type s
$l.cap_j$	Lower bound of capacity available in hospital j
$u.cap_j$	Upper bound of capacity available in hospital j
$l.op_s$	Minimum number of operations of type s to be performed
$u.op_s$	Maximum number of operations of type s to be performed

bounds on the capacity available for performing operations ($l.cap_j$ and $u.cap_j$), and the minimum and maximum number of operations to be performed ($l.op_s$ and $u.op_s$). The complete notation is summarized in Table 3.1.

The mathematical model representing the allocation problem is the following:

$$
\begin{aligned}
& min \quad M && (3.1)\\
& \text{s.t.} \sum_{j,s} c_s x_{js} \leq b && (3.2)\\
& \sum_{j} x_{js} \leq d_s \qquad \forall s && (3.3)\\
& l.cap_j \leq \sum_{s} cap.op_s \cdot x_{js} \leq u.cap_j \quad \forall j && (3.4)\\
& l.op_s \leq \sum_{j} x_{js} \leq u.op_s \quad \forall s && (3.5)\\
& x_{js} \in R^+ \qquad \forall j,s
\end{aligned}
$$

Function (3.1) is the objective function and it will be discussed in the following. Constraint (3.2) simply states that the total cost related to the planned volume of operations cannot exceed the available monetary resources. Constraints (3.3) bound the number of performed operations not to exceed the population requests. The volume of operations allocated to the hospitals is subjected to capacity constraints (3.4). Since surgical specialties differ in terms of human resources, time and space needed for interventions, the coefficient $cap.op_s$ refers to the capacity requirements of the specific type of intervention s. If we give a null value to $l.cap_j$ and infinite value to $u.cap_j$, we are assuming the extreme possibilities to close and/or open hospitals. Finally, ethic or epidemiological reasons can impose to treat a minimum number of patients, or not to exceed a maximum limit, for different types of operations, as represented by thresholds $l.op_s$ and $u.op_s$ in constraints (3.5). These constraints are relevant due to the alarming problem of over-treatment, which has been spreading during the last decade [93].

When the commissioner's perspective is taken into account, the objective function has to represent the maximization of patients' health conditions, and this objective can be expressed in several ways. In this thesis, we considered four different expressions related to mortality.

The first objective function represents the minimization of the total mortality, defined as the total number of patients who did not survive to 30 days after surgery:

$$\min \quad M_1 = \sum_{j,s} x_{js} \cdot m_s(x_{js}). \qquad (3.6)$$

A possible variation of the first objective function consists in adding a term for those patients who die because not operated, as shown in (3.7):

$$\min \quad M_1 = \sum_{j,s} x_{js} \cdot m_s(x_{js}) + \sum_{s} [(d_s - \sum_{i} x_{js})^+ \cdot m.bis_s]. \tag{3.7}$$

The first term did not change, while the second one now represents the group of people who requested to have surgery, but did not receive any treatment within the considered hospitals. Parameter $m.bis_s$ can have a double interpretation: it can be seen as the mortality risk of these patients who are to be redirected to other structures; or it can correspond to a penalty that is given to the defaulting hospitals. Since discovering the size of this specific group of patients requires information not available to us, in the following we will refer to the first formulation of M_1 (as in equation (3.6)).

The second objective function aims at reducing the average mortality:

$$\min \quad M_2 = \frac{1}{JS} \sum_{j,s} m_s(x_{js}). \tag{3.8}$$

While the first objective function focuses on the overall quality of healthcare services in all the territory, function M_2 promotes a fair attention for people living in different areas as it takes into account the average mortality risk of each kind of operation within the considered hospitals.

Focusing on the average, however, does not necessarily lead towards an equal level of mortality. For this reason, we explicitly considered the minimization of the mortality differences among the hospitals. In particular, the third objective function heads towards the minimization of the sum of mortality risk variances across interventions. Let μ_s be the average mortality risk of operation s among all the hospitals, i.e.,:

$$\mu_s = \frac{1}{J} \sum_{j} m_s(x_{js}). \tag{3.9}$$

The variance of operation s among all the hospitals is given by

$$\frac{1}{J}\sum_{j}[\mu_s - m_s(x_{js})]^2. \tag{3.10}$$

Considering all the operation types, function M_3 can be formulated as:

$$\min \quad M_3 = \sum_{s}\frac{1}{J}\sum_{j}[\mu_s - m_s(x_{js})]^2. \tag{3.11}$$

Function M_3 considers the homogeneity of service quality, both among hospitals and among interventions. In fact, the variance of the mortality risk is considered for each intervention. On one side, by being the sum of all the variances, rather than the average on the intervention types, M_3 guarantees that all the intervention variances are in the spotlight. On the other side, it neglects different weights for variances of different magnitudes.

Finally, the fourth objective function focuses on the minimization of the average range of the mortality, which represents the minimization of disparity:

$$\min \quad M_4 = \frac{1}{S}\sum_{s}[\max_{j}(m_s(x_{js})) - \min_{j}(m_s(x_{js}))]. \tag{3.12}$$

It should be noticed that, apart from M_1, all the other objective functions do not include, in their mathematical formulation, weights related to operation volumes. Hence, by placing the same emphasis on all the operations outcomes, they allow operation types with small overall volumes to have the same importance as operation types with high overall volumes. However, the ultimate aim of this homogeneity is to avoid any sort of inequality.

As it will be shown in section 3.3, each objective function raises awareness in a specific aspect of the health policy objectives, and, at the same time, increases the risk of hindering some health priorities for the population.

3.3 Health policy objectives

The basic functional form representing the volume–outcome association is a monotonic non-increasing function. Several studies have analyzed this relation using real data. However, a crucial point is represented by their methodological shortcomings [8]. In fact, the validity of the studies can be compromised if they do not take into account the higher likelihood for large hospitals of receiving more complex cases. To avoid this problem, proper risk adjustment techniques can be used to extensively rescue the confounding danger.

For this reason, in our analysis, we did not use the rough data on mortality, but we took into account only the mortality curves provided by the Italian PNE study, the analytical and statistical precision of which have been documented [10]. In fact, throughout the analysis of national data, potential confounders, including the age and the presence of co-morbidities, have been considered. Furthermore, the PNE developed a mathematical model in order to identify the best functional forms fitting the data on outcomes and volumes [1].

To have some first insights on the solution structure promoted by the different objective functions, we considered the simplest allocation case, with only two hospitals (i.e., $J = 2$) and when facility closure is not allowed (to this purpose, we set $l.cap_j = 10$, and $u.cap_j = d_s - 10$). Furthermore, we did not take into consideration the cost constraint (5.2). We experimented with a total demand d_s equal to 300.

We focused on a unique surgery type (i.e., $S = 1$), which is the hip fracture intervention, whose volume–outcome association is shown in Figure 3.1. It is a clear example of procedure that exhibits a falling mortality risk in the measured range, tending to flatten out at a volume higher than 200.
As for $l.op_s$ and $u.op_s$, we set both to the number of interventions required d_s, i.e., we have imposed to satisfy the total demand and to avoid over-treatment.
Figure 3.2 shows the behavior of the first objective function, M_1. The difference in the mortality resulting from different volume allocation is not very large. However, more insights can be obtained from Figure 3.3, in which function M_1 is reported on a reduced vertical scale. It can be noticed that the total mortality reaches its maximum in case of balanced solutions (150 interventions per hospital), while the most extreme configurations (10 volume to hospital 1, and 290 volume to hospital 2, or vice versa) correspond to the lowest total mortality. The difference between the highest and the

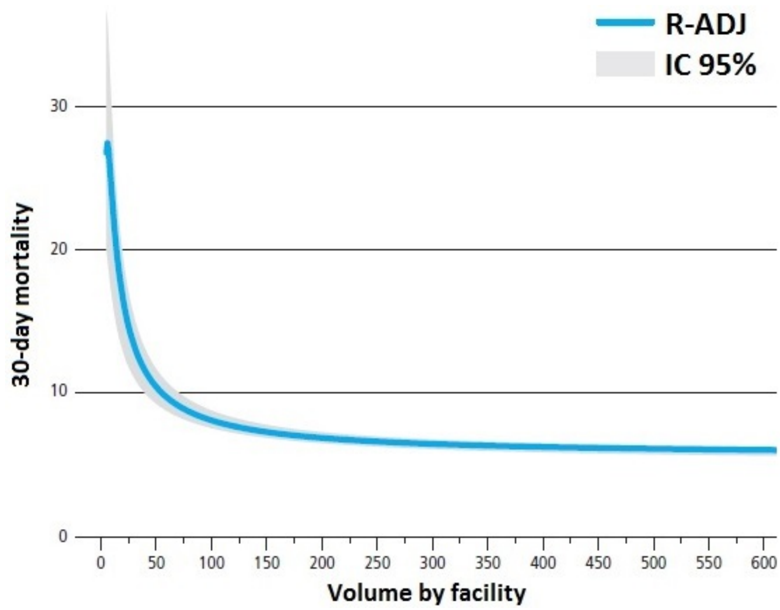


Fig. 3.1 Hip fracture; association between 30-day mortality and volume of activity by facility, Italy 2011 [1]

lowest mortality is only 0.7. However, this saving represents around the 3% of the total deaths, which is not a negligible percentage, in general.

Figure 3.3(b) reports a contour plot, which allows to gather a deeper understanding about the M_1 global behavior. The plot reports on the x and y axis the volumes allocated to the two hospitals. Hence, each point on the plot represents a configuration. Considering the continuous lines, each line corresponds to a specific M_1 value. In particular, given a line, each point belonging to it corresponds to a volume allocation that leads to the M_1 value associated to the line. Notice that the total volumes corresponding to the points on the same continuous line are different. The oblique dotted lines, instead, show different levels of demand, i.e., points of a given line have a constant sum of the x and y coordinates. However, points of a given line correspond to different M_1 values. Through the analysis of the contour plot, it appears that M_1 always fosters unequal allocations among hospitals.

The solutions suggested by objective function M_1 drive towards a global better outcome. However, when considering each hospital separately, we find out a completely uneven outcome distribution. In fact, when one of the two facilities remains open with a low volume, disparity between their outcomes becomes relevant. In the case of the hip fracture surgery, a volume of 10 interventions corresponds to

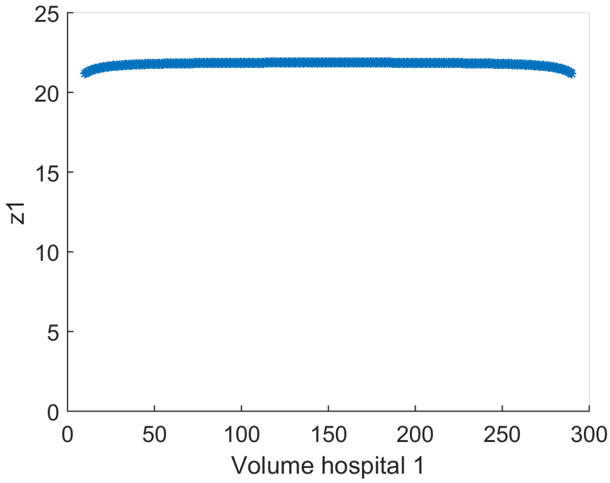


Fig. 3.2 Function M_1 behavior in hip fracture surgery, with total demand equal to 300

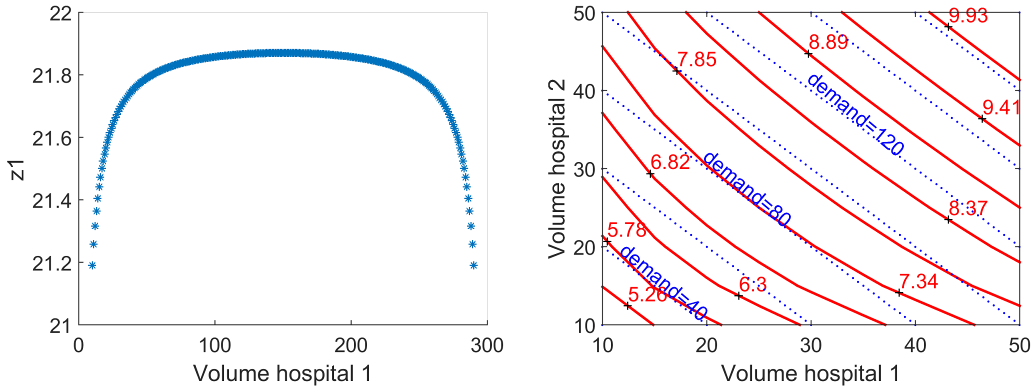
a mortality risk of 23.7%, whereas a volume of 290 procedures corresponds to a mortality risk of 6.5%.

Differently from function M_1, minimizing the average mortality (function M_2) aims at responding to the equity need in the population. If we observe the behavior of the average mortality (Figure 3.4(a)), it can be noticed that M_2 favors completely balanced solutions, even though it does not take into account global quality in terms of total mortality levels. In fact, M_1 ensures to the majority of people an excellent quality level, but it penalizes a minority of patients. On the contrary, M_2 does not heavily penalize a group of patients, but it does not foster excellence either.

These results, however, depend on the functional form of the mortality curve. In particular, if the volume–outcome association was represented by a linear shape, function M_2 would give birth to the solutions shown in Figure 3.4(b). The reason of such a behavior is that M_2 takes only the average condition into account, neglecting the single hospital performance. In other words, the objective function developed for equity may endanger equity itself.

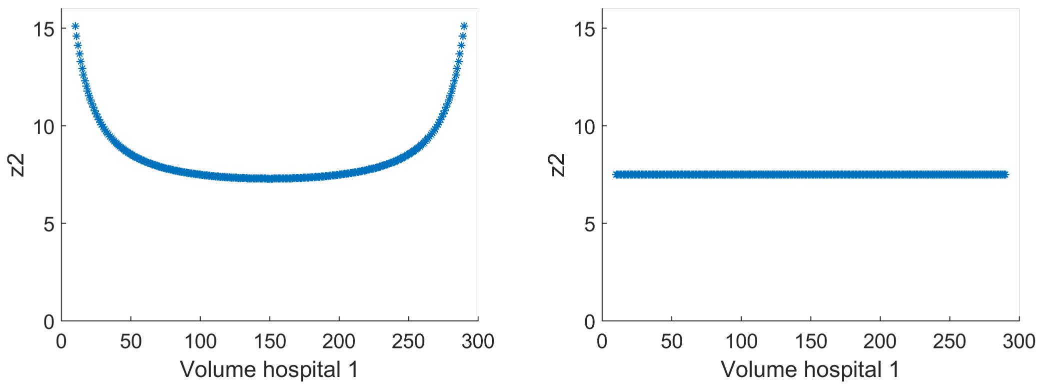
The third and fourth objective functions, i.e., M_3 and M_4, overcome the problems connected to the mortality curves functional forms. Figure 3.5 shows their behavior with both linear mortality curve and PNE mortality curve.

It can be noticed that they support equity in both cases, although attention must be paid when the number of facilities increases. In fact, using the range (i.e., M_4)



(a) (b)

Fig. 3.3 Function M_1 behavior in hip fracture surgery; total demand equal to 300 (a), multiple total demands (b)

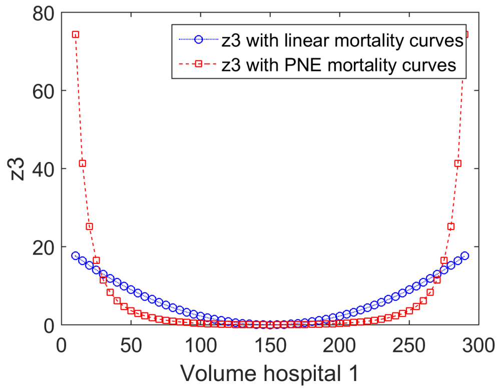


(a) (b)

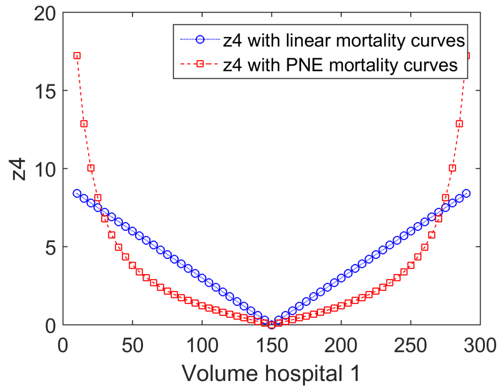
Fig. 3.4 M_2 behavior in hip fracture surgery (total demand equal to 300); PNE mortality curve case (a) and monotonic linear curve case (b)

could expose to the risk of decreasing both equity and quality. As an example, consider a case with three hospitals and assume that the three configurations reported in Table 3.2, i.e., A, B and C, are available. It can be observed that the configurations with the same maximum and minimum mortality risks would be evaluated with equal scores, even though one of them has much more dispersion in mortality risks (i.e., configuration B) or globally higher mortality risks (i.e., configuration C). Despite their differences, all the configurations result in $M_4 = 15\%$.

From the performed analysis, it can be noticed that, although the volume–outcome association can be effectively used to support the allocation of the op-



(a)



(b)

Fig. 3.5 M_3 (a) and M_4 (b) objective function behavior in hip fracture surgery (total demand equal to 300)

Table 3.2 Example of configurations to be evaluated with the average range objective function (M_4)

	Hospitals mortality risks		
Configuration	Hosp1	Hosp2	Hosp3
A	16%	1%	5%
B	16%	1%	15%
C	40%	35%	25%

erations among the healthcare facilities, the different objective functions could lead to different solutions in terms of quality and equity of the healthcare service. Each objective function, which acts as an indicator, by focusing on one aspect, neglects some others. Therefore, policy makers should be careful in evaluating the existing trade–off behind each of them, and ultimately choosing which objective to pursue.

3.4 Case studies

In this section, we show the application of the volume–outcome association on four case studies considering the reconfiguration of Piedmont Region surgery wards. Differently from the analysis performed in the previous sections, the case studies take into account real life factors such as the reallocation of interventions, the geographical distribution, the specialization thresholds and the performance of the hospitals, respectively.

In order to focus specifically on the volume–outcome association, we did not take into consideration the cost constraint (3.2), and, as discussed before, we assumed a deterministic mortality curve, i.e, we did not consider the effect of the uncertainty. Moreover, since we are interested in exploring all the possible territorial configurations, we admitted hospitals closure, which means setting the hospital capacity lower bounds to zero ($l.cap_j = 0$). As for $l.op_s$ and $u.op_s$, we set both to the required number of interventions, d_s, thus imposing to satisfy the total demand and to avoid over-treatment.

Eventually, we compared our solutions to the actual configuration, using the real data available from the PNE report [2].

The optimization problem has been solved using MATLAB *fmincon* function with the *Sequential Quadratic Programming* method [94].

3.4.1 Reallocation of interventions

In the first case study, we considered the reconfiguration of a group of surgery wards in an Italian Region, Piedmont. We took into account five types of surgical interventions for which there is evidence of the volume–outcome association: aortocoronary bypass, surgery for colon cancer, surgery for pancreatic cancer, surgery for lung cancer, nonruptured abdominal aortic aneurysm (aaa). We considered three hospitals that have performed those surgeries in 2013, all of them located in Turin (the biggest city in the Region), thus involving a provincial reconfiguration.

We have tested two scenarios. The first scenario deals with an allocation made from greenfield. We did not insert any constraints, apart from the hospital total capacity. The solution obtained using M_1 clearly concentrates the interventions in single structures. Analyzing the differences between proposed and actual provincial configurations in terms of operation volumes, it clearly appears the suggested specialization of hospitals in specific interventions, rather than an apparently unjustified division of interventions among structures.

The second scenario has an additional constraint concerning the adaptability of surgeons and operating rooms, which better (although not exhaustively) represents reality. Hospitals keep on performing the types of operation they have performed the previous year. In fact, surgeons usually specialize in one kind of interventions and hospitals too as a consequence. In this scenario, M_1 still drives to specialization, especially for the riskier interventions. However, the solution leads to a higher total mortality (as compared with the first scenario), because of the impossibility to freely allocate operations. Hence, as expected, imposed constraints prevent from optimal solutions.

Figure 3.6 shows the three configurations (actual, proposed from greenfield, proposed with specialization constraints) in terms of volume allocation between hospitals, while Figure 3.7 reports the related calculated total mortality (i.e., M_1) and average mortality (i.e., M_2) values of each type of intervention. The proposed

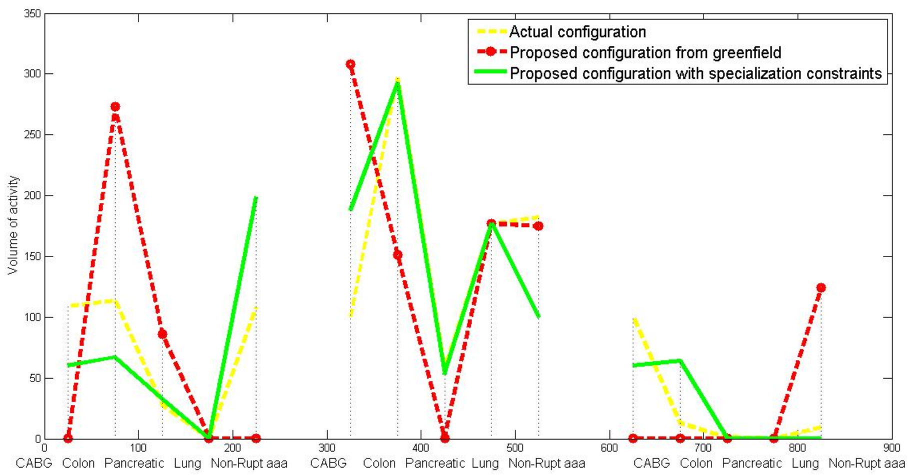


Fig. 3.6 Volume allocation between hospitals in different configurations

configuration from greenfield drives to the lowest values of total mortality and average mortality.

Interestingly, the configuration with the best outcome requires several changes in the territorial organization, since hospitals need to re-arrange their wards in order to concentrate on some operations (not necessarily performed in the hospital before the reconfiguration), neglecting others (i.e., closing the related wards). It is also noteworthy the worse outcome caused by the proposed configuration with specialization constraints.

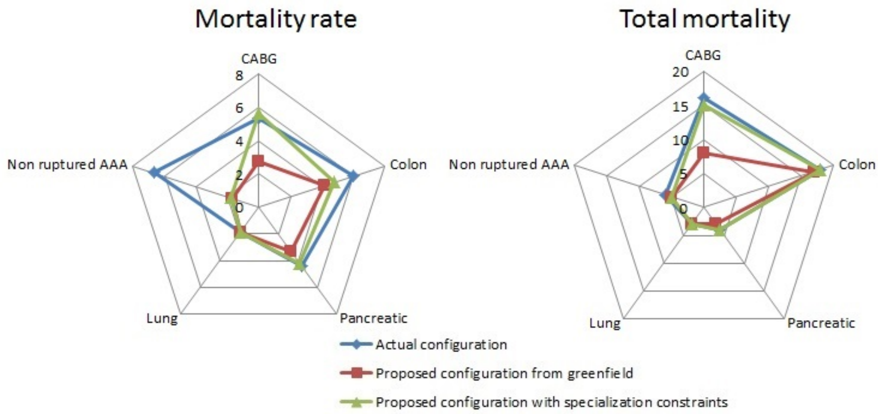


Fig. 3.7 Comparison between average mortality (M_2) and total mortality (M_1) of each type of intervention in the three configurations

The regional consequences are not to be ignored, both for hospitals that would host more patients and hospitals that would be forced to close. Also, concentration could lead to unfair outcomes in different areas of the territory.

3.4.2 Geographical distribution

In the second case study, we investigated the geographical distribution of hospital structures. The aim was to examine how a total amount of operations can be assigned to the hospital structures in a territory (defined in terms of geographical boundaries), taking into account their geographical position. The geographical perspective is especially worth of interest because of its policy implications. In fact, in healthcare systems that have an organization similar to the Italian one, territorial decision makers do have the opportunity to choose which hospitals can/must perform a certain type of intervention.

As a territory, we again focused on Piedmont. Piedmont Region is composed by eight provinces, of different size and population density. However, in order to the get the first insights on the territorial configuration problem, we did not take into account the population dispersion in the territory and its density. The main assumption in our analysis is that each province takes care of its population, regardless of its size.

We considered one type of surgery, bladder cancer intervention, which has been performed in 2013 by several facilities scattered in the regional territory, as shown in Figure 3.8. In 2013, in Piedmont, the 84% of bladder cancer interventions have been elective cases [95], thus the majority of the people actually had the possibility to book date and place of their intervention in hospitals wards. The corresponding mortality

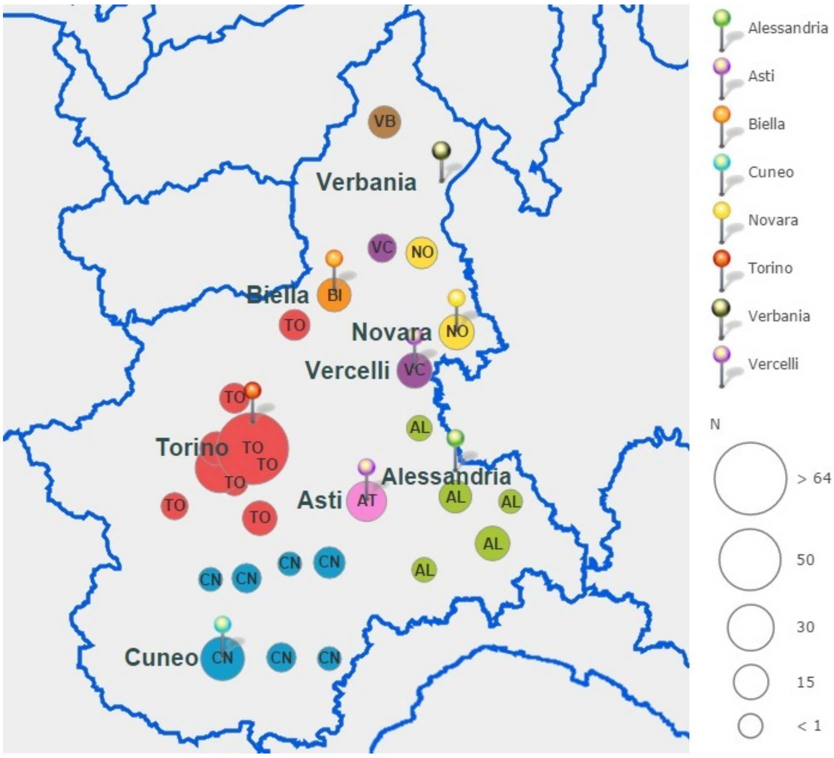


Fig. 3.8 Case study 1. Geographical distribution of volumes for bladder cancer interventions performed in Piedmont in 2013 [2]

curve for bladder cancer intervention developed by PNE is shown in Figure 3.10. There is a clear descending trend: for volumes smaller than 20 (which are performed in Italy by 80% of the facilities) the mortality risk can vary between 20% and 6%, while for volumes larger than 20 the mortality risk varies only between 6% and 1%.

In the majority of the provinces there are hospitals that are closely located, and that perform an extremely small number of interventions. In particular, the distribution of volume by facility and by province can be observed in Figure 3.9(a). Each rectangle corresponds to the volume performed by a facility, and they are categorized by province. In Figure 3.9(b), an alarming concentration of hospitals that perform a number of interventions corresponding to the highest mortality risk (assuming that their outcomes correspond to the ones predicted by the mortality curve) can be observed. In fact, the 25% of the hospitals in Piedmont has a number

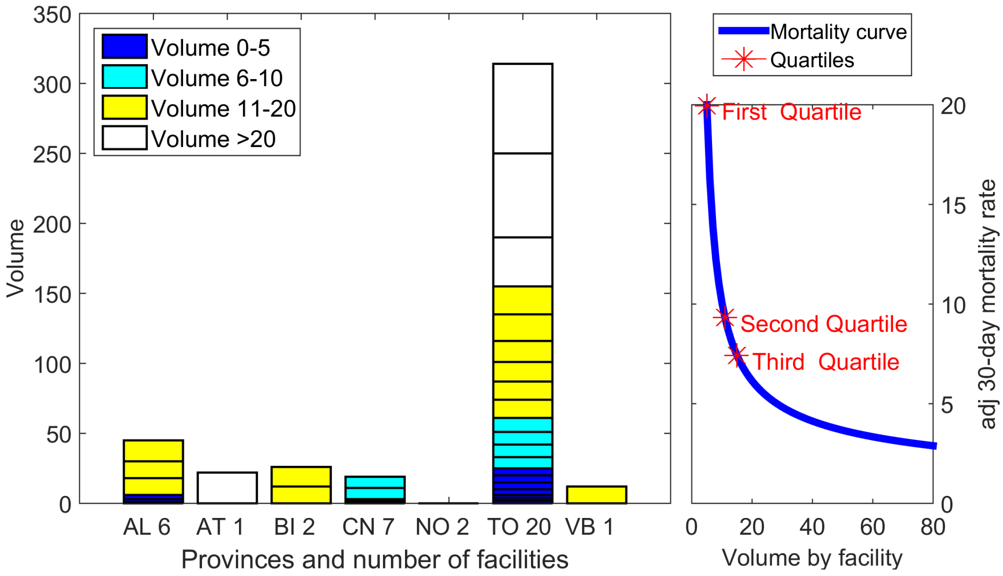


Fig. 3.9 (a) (b)
Case study 1. Distribution of volumes for bladder cancer interventions performed in Piedmont provinces in 2013 [2]

of interventions smaller than 5, and only the 12% of the facilities perform more than 20 operations.

To take the geographical issues into account, we added to the model presented in Section 3.2 constraints imposing that, for each province, at least one hospital must operate. They represent our assumption that each province has to take completely care of its population.

As expected, the optimal allocation (dots and quartiles in Figure 3.10) is given by the complete concentration of intervention in a unique hospital per province, which is completely opposed to the real actual configuration. It can be noticed that now only the 25% of facilities (first quartile in Figure 3.10) performs less than 20 interventions per year. When considering the related mortality, a clear increase of better outcomes, if compared to the real ones, can be observed. The optimal solution decreases the total mortality from 45 to 15. This remarkable difference is due to the closure of all the hospital wards that, making a really small number of interventions, were unfortunately having bad outcomes. The intervention concentration guarantees a lower mortality risk for the unique active hospital and thus a lower total mortality.

Obviously, the minimum number of required hospitals per province does have an impact on the decrease in the total mortality achievable by volume reallocation.

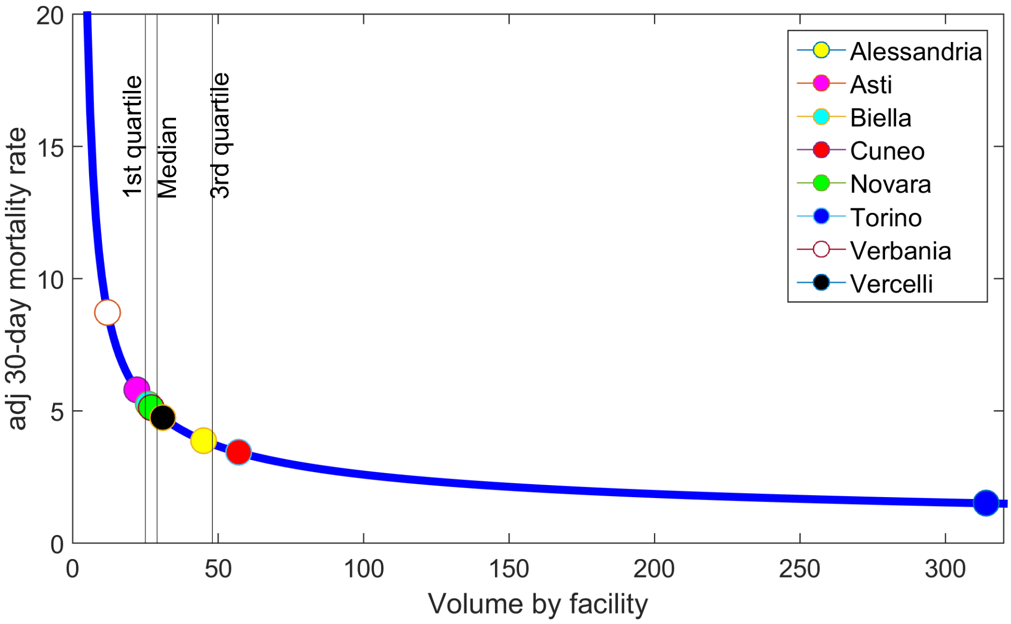


Fig. 3.10 Case study 1. Bladder cancer, analysis of the association between 30-day mortality and volume of activity by facility, Italy 2011 [1]

However, since patients may perceive mobility as a detrimental aspect [96], policy makers could increase the minimum number of required hospitals per province, in order to decrease the distance that patients have to travel. By increasing the minimum number of required hospitals, we obtain the expected result that the higher the minimum number, the worse the outcome. In fact, the presence of several hospitals per province increases the possibility to have wards performing a small number of interventions, which negatively impacts on the global outcome level. Figure 3.11 shows the optimal value of function M_1 and M_2 for different imposed minimum numbers of hospitals per province. The reduced vertical scale helps the reader to appreciate the change in values assumed by the objective functions.

All in all, results show the impact that the territorial dispersion has on hospital outcomes. The distribution of mortality risks among hospitals given by the optimal configuration, which favours the concentration of interventions in a unique hospital per province, guarantees global better clinical outcomes, and ensures patients from each province to have a hospital nearby.

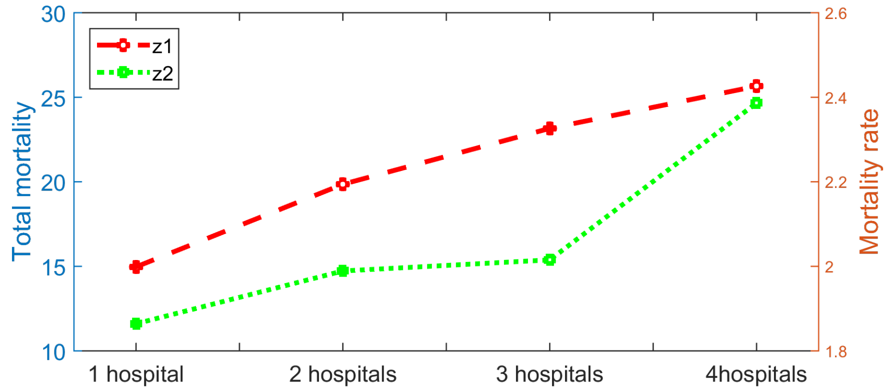


Fig. 3.11 Case study 1. Total mortality and average mortality for different minimum hospitals imposed per province

3.4.3 Specialization thresholds

In the third case study, we considered the hospital specialization, that is, the opportunity for the hospitals of focusing on fewer types of interventions and ,hence, performing higher volumes of them. In fact, based on the volume–outcome association, and as shown in the previous case studies, the lowest mortality risk corresponds to the situation in which each hospital in a territory performs a single type of surgery. However, in the real case, the demand is scattered among facilities.

Although a variety of reasons support such dispersion, the volume–outcome association shows a substantial steep increase for mortality risks corresponding to extremely small volumes of activity. As a consequence, as mentioned in Section 2.3, it has been recently advocated the existence of a threshold on the volume performed per facility. This request comes from several research groups (e.g., the Leapfrog group [88]) and it has been embraced by policy decision makers. For example, in 2015, the Italian board of health has imposed a minimum threshold, in terms of volumes of activity, for six types of interventions [97].

The aim of this third case study is to investigate the rationale behind the threshold existence and to analyze the impact of threshold values.

We considered two types of surgery: prostate cancer and colon cancer. This choice is motivated by the fact that the PNE mortality curves for these two surgeries have been developed also for a very small number of interventions, which is the case we are interested in. In Figure 3.12 we can notice that, for a number of interventions smaller than 10, the prostate cancer shows the highest mortality risk while the colon

cancer is slightly less steeper.
In 2014, in Piedmont Region, more than 72% of colon cancer interventions and

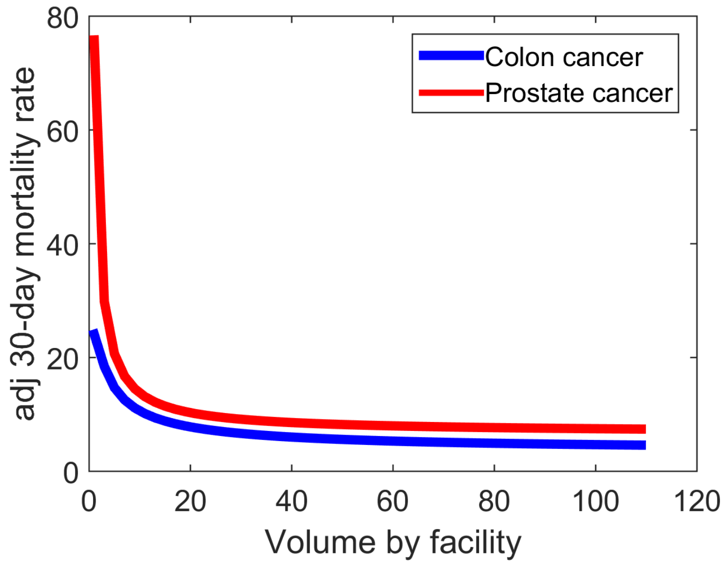


Fig. 3.12 Prostate cancer and colon cancer; analysis of the association between 30-day mortality and volume of activity by facility, Italy 2011[1]

more than 99% of prostate cancer interventions have been elective [98]. Hence, the decision makers could influence that high percentage of elective patients by steering them towards specific facilities.

In the Piedmont Region, we considered the province of Alessandria that covers 3558.83 km^2 and has eight hospitals performing the mentioned interventions. The considerable dispersion in the number of interventions currently allocated to each hospital (reported in Table 3.3) causes the presence of volumes smaller than 10 and thus high mortality risks.

We assumed each hospital to have sufficient capacity to perform the total number of prostate cancer and colon cancer interventions actually performed in 2013.

The threshold on the number of interventions works as follows. For each facility, the number of interventions performed in the past year is compared to the threshold. If the number of interventions is higher than the threshold, the hospital is constrained to remain open and to perform at least a number of interventions equal to the threshold. On the contrary, if the hospital has performed a number of interventions that is lower than the threshold, the specific ward has to be closed, that is, it is

Table 3.3 Case study 2. Volumes of colon and prostate cancer interventions performed in Alessandria province's hospitals in 2013 [2]

	Volumes by facility	
Hospital	Colon cancer	Prostate cancer
Hosp 1	102	51
Hosp 2	1	41
Hosp 3	1	0
Hosp 4	12	0
Hosp 5	29	11
Hosp 6	49	28
Hosp 7	31	14
Hosp 8	25	0

removed from the optimization model. In this case, the volume of interventions performed in the previous year is redistributed among the other wards.

We experimented with threshold values from 0 (no threshold) to 40. Figure 3.13 reports the results from the case of threshold equal to 30, showing the comparison between the real volume allocation in 2013 and the optimal volume allocation. It can be noticed that the combined effect of thresholds and total mortality minimization leads to a configuration in which most of the volume is allocated to a unique hospital (called *big facility*). All the other open hospitals perform only a number of interventions equal to the threshold (*small facilities*), which still is not associated with a good outcome.

This observation leads to the expected result that the higher the threshold, the more all the *small facilities* will guarantee a better outcome, even if the majority of the volume will still be allocated to the *big facility*. From the analysis of all the thresholds from 0 to 40, it can be observed an improvement in the objective functions (Figure 3.14). Specifically, M_2, M_3 and M_4 have a severe decrease, thus confirming that higher thresholds guarantee a better average performance and more equity. As for the total mortality M_1, its trend is not as consistent as the other ones. In particular, the reduced vertical scale highlights that the prostate cancer has intervals that are steady instead of decreasing, whereas the colon cancer has increasing tendency for several volume intervals followed by sudden falls (*jumps*). A deeper analysis of the results showed that a jump occurs when the threshold is high enough to force one additional hospital ward to be closed. When this happens, the volume

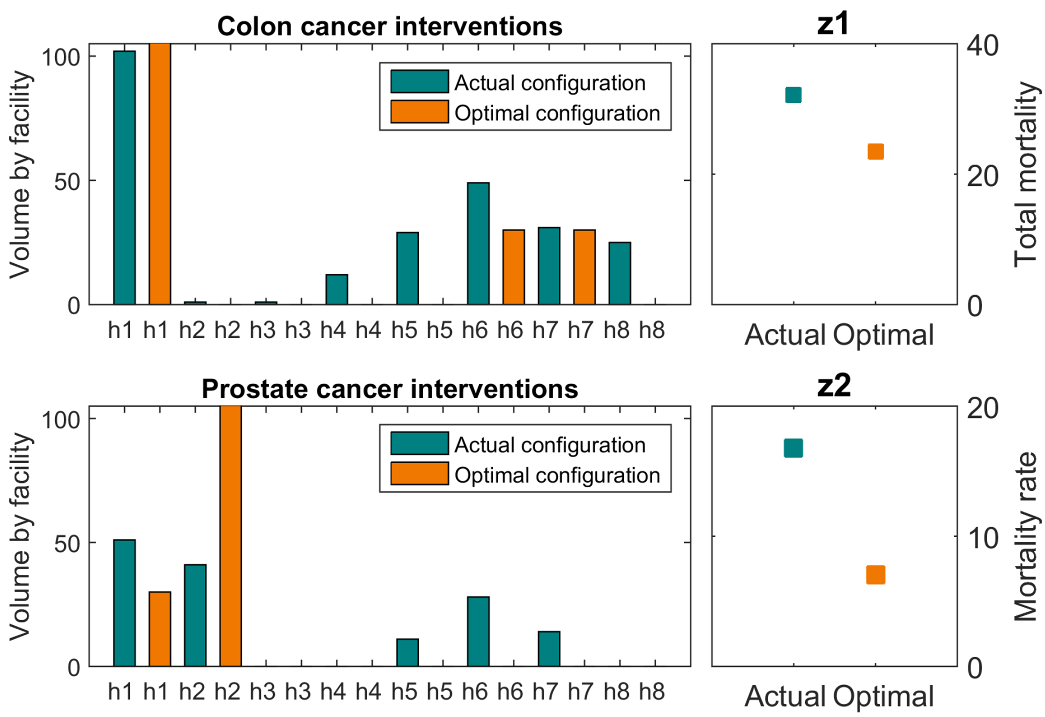


Fig. 3.13 Case study 2. Comparison among optimal allocation of volumes (threshold 30) and volumes actually performed in 2013, Alessandria province

of the hospital under-threshold is allocated to the *big facility*, thus guaranteeing a remarkable improvement in the outcome. From that point on, as the threshold increases, the objective function value increases as well. This is due to the fact that each additional unit increase in the threshold forces the *small facilities* to perform each one more intervention, and thus the *big facility* to loose some of its interventions. This happens until the threshold is high enough to originate an additional hospital closure, causing the M_1 function value to have a sudden fall, and so on.

Results from this case enlighten the benefits of forcing that only hospitals with expert surgeons (i.e., those who have performed high volumes in the past) can perform the interventions. Nonetheless, no clear cut-off point exists, being the higher the threshold, the better the outcome.

3.4.4 Hospital performance

In the forth case study, we focused on hospital performance. Although the volume–outcome association assigns to each volume a unique outcome, considering real data

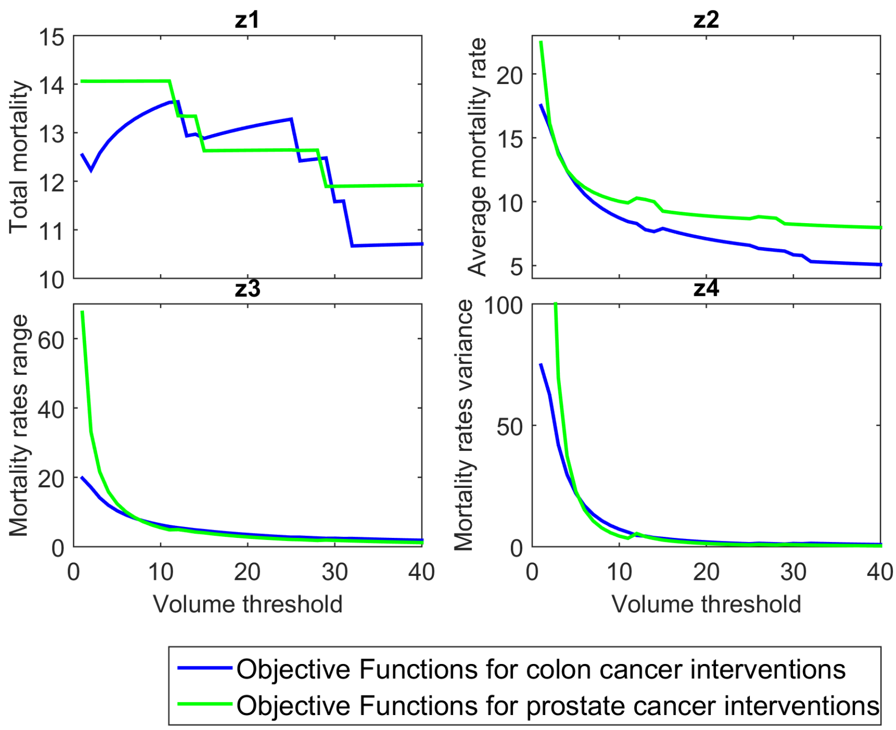
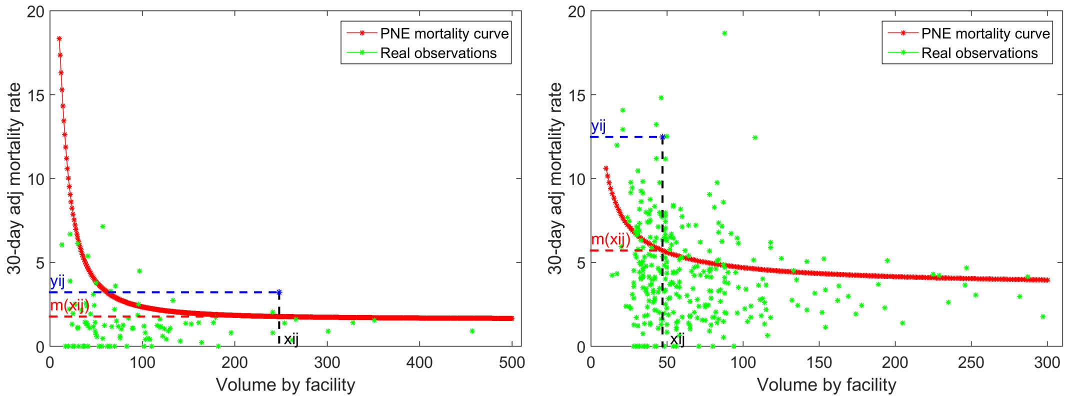


Fig. 3.14 Case study 2. Comparison among objective functions' values with thresholds 0-40

of adjusted mortality risks, differences in the outcomes for the hospital executing the same number of interventions can be noticed. As an example, Figure 3.15 shows for colon and lung cancer interventions both the volume–outcome association (the red curve) and the real risk-adjusted observations from Italian hospitals (green dots) [1].

As discussed in section 3.1, surgeon's talent and general uncertainty cause the real risk-adjusted mortality to differ from the value predicted by the mortality curve. In this case study, we neglected the general uncertainty and only focused on the variability caused by physicians' ability. Moreover, we assumed that the variability on the mortality generated by physicians' excellence/incompetence is controllable. This assumption can be reasonable since the improvements in surgeons' skills could be guided by specific policies.

To consider the differences in the mortality only due to surgeons' talent, we assigned to each hospital a performance coefficient based on the comparison between the real outcome of the hospital (real data are reported in [3]) and the outcome suggested by the mortality curve [1]. These two items are comparable since both of them consider risk-adjusted data.



(a) (b)

Fig. 3.15 Effective risk-adjusted outcomes (2013) compared to PNE mortality curves, for lung (a) and colon (b) cancer interventions

Let $m.real_{js}$ be the effective mortality risk of hospital j for intervention s, and $m_s(x_{js})$ the mortality risk estimated by the mortality curve when the number of interventions is equal to x_{js}.

For each volume x_s, we considered the absolute value of the deviation (i.e., the difference between observed and predicted values) for each facility that is performing that volume, i.e., for which $x_{js} = x_s$, and calculate the average absolute deviation as follows:

$$dev(x_s) = \frac{1}{n(x_s)} \sum_j \left| m.real_{js} - m_s(x_{js}) \right| \tag{3.13}$$

where $n(x_s)$ is the total number of hospitals that are performing volume x_s. These hospitals are represented by all the points in Figure 3.15 that have the same x coordinate.

Once the average deviation $dev(x_s)$ has been computed for each volume x_s, we fitted these values, obtaining the $DEV(x_s)$ curves shown in Figure 3.16.

The DEV curve was used to create a *deviation interval* around the mortality curve (i.e., the volume–outcome association). The deviation interval should not be confused with the confidence interval, since the deviation interval was simply determined by adding to, and subtracting from, the mortality curve the DEV curve (Figure 3.17). By using the deviation interval as a bound for the variability of the real data, we are not considering the variability of the real data that are very far from the predicted values (the outliers). Since outliers are likely due to uncontrollable and

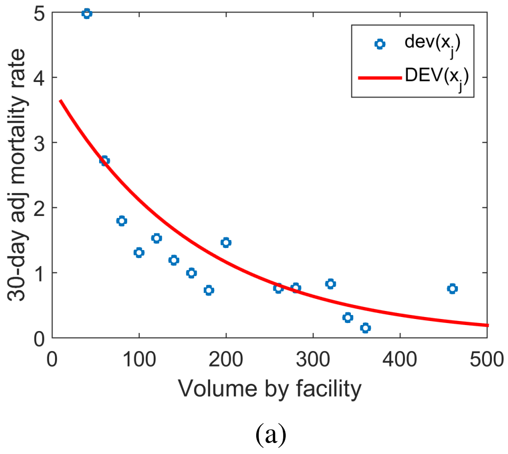


Fig. 3.16 Interpolation of the average absolute value deviations, for lung (a) and colon (b) cancer interventions

accidental factors, neglecting a huge part of their variability corresponds to consider only the variability caused by surgeons' talent, rather than the whole variability determined also by uncontrollable factors (i.e., general uncertainty).

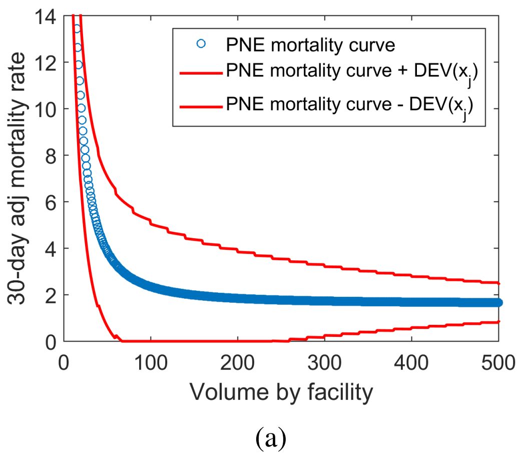


Fig. 3.17 PNE mortality curves with deviation bounds, for lung (a) and colon (b) cancer interventions

Eventually, we defined the performance coefficient v_{js} of hospital j for surgery s as:

$$v_{js} = \operatorname{sgn}\left(m.real_{js} - m_s(x_{js})\right) * \min\left(1, \frac{\left|m.real_{js} - m_s(x_{js})\right|}{DEV(x_s)}\right). \tag{3.14}$$

The coefficient v_{js} assumes values from -1 to +1. Positive values correspond to mortality risks higher than the predicted ones, hence to poor performances. On the contrary, negative values represent excellence. These coefficients were used in the allocation of volumes, where we assumed that the hospital performance in the future will be the same as that achieved in the past.

Differently from the three previous case studies, where constraints were added to the basic model of Section 3.2 to include real life factors (e.g., specialization thresholds), in this case we modified the objective function. In particular, M_1 has been modified as follows:

$$min \quad z_{policy} = \sum_{j=1}^{J} x_{j_s} * m_s(x_{js}) + \sum_{j=1}^{J} x_{j_s} * v_{js} * dev(x_{js}) \tag{3.15}$$

where the first term is the volume of activity multiplied for the mortality risk of the hospital (i.e., the original M_1), while the second term has the role of increasing/decreasing the mortality depending on the coefficient v_{js}.

To evaluate the impact of the hospital performance on allocation, we considered three hospitals from the province of Turin (Piedmont region). It should be noticed that, in Piedmont Region, in 2014, 72% of colon cancer interventions and 97% of lung cancer interventions have been elective, hence there exists the possibility for health policies influencing patients' choice of hospital. Figure 3.18 reports the three hospital outcomes (green dots) compared to the mortality curves of colon and lung cancer interventions, and their performance coefficients. All the hospitals performed better than how the mortality curve would predict, even though there are differences among them. In particular, policy makers could leverage the better performance observed in one procedure by fostering the facility specialization in that specific surgery.

The actual allocation of the interventions is shown in Figure 3.19(a). Comparing Figure 3.18 and Figure 3.19(a), it can be noticed that the actual volume allocation does not completely comply with the hospitals competences. For example, the Don Bosco Hospital has its best performance coefficient on lung cancer intervention, but it uses the majority of its capacity in colon cancer interventions.

Considering the optimal configuration (Figure 3.19)(b), it appears that each hospital specializes in its most advanced skills. For instance, the Don Bosco Hospital

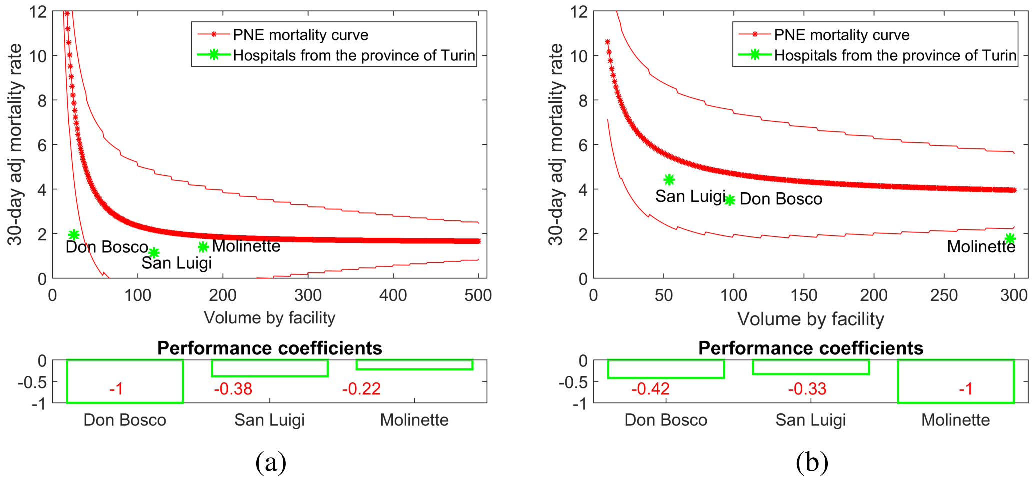


Fig. 3.18 Case study 3. Actual outcomes compared to mortality curves, and performance coefficients for three hospitals of the province of Turin, for lung (a) and colon (b) cancer interventions

is suggested to perform only lung cancer interventions, while the colon cancer interventions are concentrated in the Molinette hospital, which had the highest performance coefficient.

The comparison among the actual and the optimal configuration (Figure 3.20) quantifies the benefit that patients gain from being treated by the most expert surgeons.

3.5 Limitations

Results from this chapter globally highlight the importance of considering the relevant factors of the real case, which allow to obtain different optimal solutions for different contexts, and confirm the paramount importance of considering the volume-outcome association within the strategic planning of healthcare services, in order to guarantee better health to the population, both in terms of quality and fairness.

Summing up, the policy maker model is totally focused on the minimization of the total mortality. Constraints are represented by threshold and capacity issue for volumes, demand satisfaction and allocation of provincial demand among provinces. Hence, the policy maker represents the institutional figure that cares about the whole

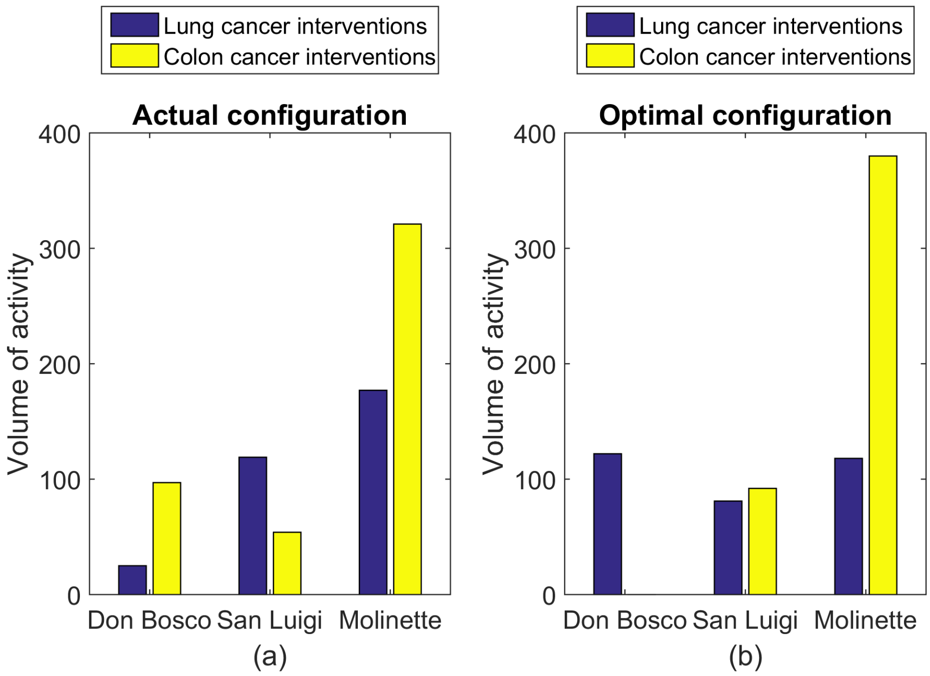


Fig. 3.19 Case study 3. Actual volumes of colon and lung cancer interventions performed in three hospitals in the province of Turin in 2012 [3] (a) and optimal allocation (b)

population health, in terms of both universal coverage and the quality of the services received. However, in the model previously reported, no attention is paid to patients' answer, i.e., patients are assumed to behave such as to confirm volumes that have been strategically planned by the policy makers.

The lack of prediction about patients' behaviours causes the actual volume distribution to be likely to differ from the one planned by the policy maker. In fact, since patients are free to choose the provider to be treated in, they could opt for travelling to any hospital in the region. The opening/closure of hospitals remain the only thing patients account for. As no hospital can forbid patients from choosing it, any change in the volume distribution among hospitals has to be managed, causing consequences on hospital capacity stress and, most importantly, on the deterioration of clinical quality. This criticality has been addressed in chapters 4 and 5.

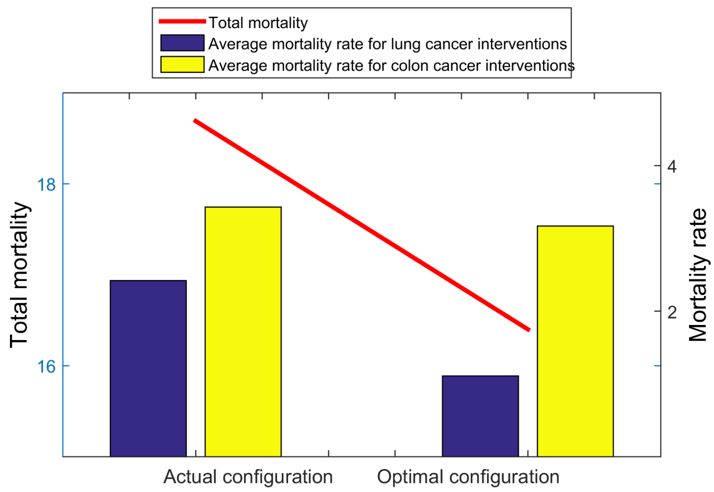


Fig. 3.20 Case study 3. Comparison among outcomes in actual and optimal configuration, for three hospitals of the province of Turin

Chapter 4

Patient perspective

This chapter focuses on patients' perspective. By changing the point of view, the problem is not any more a planning problem, centred on the volume of activity of an hospital (as considered from the policy maker's perspective), but rather it is a choice problem, which is centred on each patient that is choosing a facility. Moreover, our attention moves from the aggregated level of the hospital, i.e., volumes, to the individual level of the patient, i.e., single choices.

In healthcare systems like the Italian one (see Chapter 1), patients are free to choose the facility to be treated in. Moreover, reservations made by patients for visits or interventions cannot be refused by hospitals. Hence, even though the strategical planning is a decision that is entitled to policy makers, eventually patients have the possibility, with their choices, to confirm or completely change the regional distribution of volumes of activity.

A key issue of this research is to understand which factors can cause patients' adherence/refusal of policy makers' plans. We are then going to explore which are the main drivers of patients' choices.

4.1 Problem description

When looking at patients behaviour, two different approaches can be adopted. With the first approach, the main point of interest becomes the analysis of the reasons that motivate patients to travel, i.e., to understand how hospital characteristics (like quality) and patient characteristics (like age, clinical conditions, etc.) have an impact on the distance travelled by patients. Hence, by using a regression model, linking the dependent variable (i.e., distance) to a number of independent variables (e.g., hospital characteristics), it is possible to extract the coefficients that represent the actual behaviour shown by patients. The second approach, instead, consists in the so called *choice models*, which focus on all the elements of the choice of the hospital made by patients. This second approach is the most used in the health economics literature to study patients' preferences. In fact, the added value of these models is to observe not only what people have actually done, but also what they could have done (and eventually have chosen not to do). In terms of our research, the idea translates into observing both in which hospitals patients have chosen to be treated, and out of which alternative hospitals. The advantage of this approach consists in gathering a richer amount of information on patients' preferences, since they are related to all the alternatives, rather than to only one.

Hence, the problem analysed becomes the following one. We consider a geographic region, with its own population, and a set of hospitals (the choice set) among which the regional population can choose to be treated. We consider the population as composed of I patients denoted by i, where each patient i chooses one hospital j among J facilities, representing the choice set.

The choice of a patient for a specific hospital is considered as depending both (i) on the value perceived from receiving care in that hospital, which means, on the expected utility gained, and (ii) on the comparison among expected utilities that would be gained if seeking care from different available hospitals. It can be deducted that the decisional process of all the patients comprises a first phase of gathering information about the available hospitals, and a second phase of comparing hospitals' information. Since patients are assumed to be rational agents [66], they eventually aim at the maximisation of their utility, which means, they are going to choose the hospital that ensures them the highest expected utility.

In order to establish the value of the expected utility gained, we acknowledge as main determinants of patient utility distance and quality. In fact, in the Italian public healthcare system, patients do not pay for the service they receive [99], so monetary price does not appear among the factors affecting the decision. Hence, each hospital j can be considered for its quality attributes q_j, and its location at a distance d_{ij} from the patient i. The utility of patient i going in hospital j can thus be modelled as:

$$U_{ij} = U_{ij}(d_{ij}, q_j) + \varepsilon_{ij} \tag{4.1}$$

where ε_{ij} represents the unobserved utility, e.g., the part of patient i's evaluation of hospital j including information obtained by word of mouth and prior experience [66].

The relationships between utility and distance, and utility and quality, can be expressed as follows:

$$\frac{\partial U}{\partial d} \leq 0 \qquad\qquad \frac{\partial U}{\partial q} \geq 0 \tag{4.2}$$

These relationships imply that increasing distance reduces utility, while utility is increasing in hospital quality. Larger distance reduces utility because travelling is costly, requiring time to reach the hospital together with transport costs, both for patients and their relatives. On the contrary, as patients seek care to get better, they look for the best facility in terms of quality outcomes. Hence, patients face a trade-off when having to travel further to secure higher quality care.

Furthermore, patient perceptions of distance and hospital quality depend on their choice set of J hospitals, and their characteristics. As choice set we include all the hospitals available to patients within the geographical region at a given time. Hence, all patients in the same geographical region treated in the same year have the same choice set. Since hospitals in the choice set differ in terms of quality and distance, patients need to decide how to evaluate these two elements in order to maximise their utility.

4.2 Model formulation

As for the model formulation, choice models focus on individuals making choices out of available alternatives and consider two sets of influences on choice: individual-specific variables (i.e., related to patients) and alternative-specific variables (i.e., related to hospitals). In fact, in the literature it has been shown that the importance given to hospital characteristics (like distance and quality) in the decisional process depends also on patient characteristics, such as their age, social and clinical conditions [58, 57].

The analytical tool used has been the conditional logit random utility model (see [100] for all the analytical demonstrations). The logit assesses the probabilities of choice while the random utility model investigates patients' utility maximization. Compared to other tools, such as the multinomial and the mixed logit, the peculiarity of the conditional logit is that it concentrates only on alternative-specific variables for the utility calculation. Individual-specific variables are used only as interaction terms with the alternative-specific variables, capturing the change in the impact of alternative-specific variables due to patient characteristics. This tool fits our research interest in understanding how hospital characteristics drive patients' choice.

The utility of patient i receiving treatment in hospital j at time t is

$$U_{ij} = V_{ij} + \varepsilon_{ij}. \tag{4.3}$$

V_{ij} is the part of the utility that can be observed, including alternative-specific variables. ε_{ij} is the unobserved part, which we consider as a random error term. For each patient i, J utility formulae can be drawn to investigate the expected utility to be gained if choosing each of the available hospitals. Each patient is assumed to select hospital j from which it is possible to receive the highest utility conditional on the utilities offered by other hospital choices. The conditional random utility model, for each patient, compares utilities gained by pairs composed by the chosen alternative and an unselected alternative, and performs the pairwise comparison for all the unselected alternatives, taken one a time. In this way, it accepts the assumption of Independence of Irrelevant Alternatives (IIA), which states that the preferences among alternatives X and Y only depends on patient preferences on X and Y (and not on preferences for other alternatives).

If y_{ijt} represents the patient choice, we know for each patient i which hospital j was chosen at time t (y_{ijt} = *1*) and which was not selected (y_{ijt} = *0*). Hence, the model examines the trade-off among hospitals as revealed by the choice that each patient made. Assuming that ε_{ij} are random variables independently and identically distributed (i.i.d), the probability of patient i choosing hospital j out of a total of J

hospitals at time t is:

$$P_{ijt} = \frac{e^{V_{ijt}}}{\sum_{j' \in C} e^{V_{ij't}}}. \qquad \forall j = 1, ...J \tag{4.4}$$

The conditional logit solution process is through the maximization of the log-likelihood function:

$$lnL = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{t=1}^{T} y_{ijt} * ln\left(\frac{e^{V_{ijt}}}{\sum_{j' \in C} e^{V_{ij't}}}\right). \tag{4.5}$$

Since the function L has to be maximized, the solution process will extract parameters V_{ijt} that result in P_{ijt} being closest to 1 for the chosen option, where $y_{ijt} = 1$.

Implementation of the conditional logit requires a decision about what variables to include in the vector V_{ijt} (and thus in U_{ijt}). Our strategy has been to develop multiple models, starting from a baseline and enriching it gradually, in order to capture the added value of additional variables. Specifically, we developed four models of the patient utility function.

In the first model, we assume that utility of patient i receiving treatment in hospital j at time t is a linear additive function of distance and quality variables:

$$U_{ijt} = d_{ij} * \alpha + \sum_{k=1}^{K} q_{jk,t-1} * \beta_k + \varepsilon_{ijt}. \tag{4.6}$$

where

- d_{ij} refers to the geographic distance between the patient's place of residence and the hospital location. This hospital characteristic varies by patient and hospital but not over time.
- $q_{jk,t-1}$ is the quality measure k evaluated for hospital j at time $t-1$. For each hospital j, $\sum_{t=1}^{T} q_{jk,t-1}$ is a vector that includes the annual quality measures k for all the considered years. As in [58, 56], quality features are measured for the year before the patient choice, as we assume patients receive information with a time lag of one year. Moreover, by using the time lag, possible reverse causality between choice and hospital quality is overcome [58], since demand at time t cannot affect quality at time $t-1$.

- coefficients α and β_k capture the change in utility due to changes in, respectively, distance and quality measure k

The second model adds alternative-specific intercepts, i.e., constant terms for each hospital:

$$U_{ijt} = d_{ij} * \alpha + \sum_{k=1}^{K} q_{jk,t-1} * \beta_k + \gamma_j + \varepsilon_{ijt}. \tag{4.7}$$

As already shown in [74], γ_j acts as a hospital fixed-effect. The rational for this constant is to capture unobservable time-invariant hospital characteristics, so that these are distinguished from observable and time-varying hospital quality measures. These fixed effects might include unobservable aspects (e.g., hospital reputation), which can be correlated with observed quality measures. By inserting an alternative-specific constant, the effect of measured quality variables can be separated from the impact of constant and unobserved hospital features. In this way, model coefficients highlight if patient choice is driven by unvarying hospital characteristics (all captured by the alternative-specific intercept) or by time varying measures of quality.

The third model adds patient characteristics as interaction terms to correct coefficients for observable heterogeneity among patients. The purpose of this model is to assess whether this possible heterogeneity is of importance for the estimation of our primary variables of interest:

$$\begin{aligned} U_{ijt} = d_{ij} * \alpha + \sum_{k=1}^{K} q_{jk,t-1} * \beta_k + \sum_{l=1}^{L} x_{il} * d_{ij} * \delta_l + \\ + \sum_{k=1}^{K} \sum_{l=1}^{L} q_{j,k,t-1} * x_{il} * \lambda_{kl} + \gamma_j + \varepsilon_{ijt}. \end{aligned} \tag{4.8}$$

where

- $\sum_{l=1}^{L} x_{il}$ gathers L patient attributes, which act as interaction terms to correct coefficients for observable heterogeneity among patients;
- the coefficients δ_l and λ_{kl} express the change in the perception of, respectively, distance and quality measure k, due to the personal characteristic l.

It should be noted that equation 4.8 takes into consideration interactions between hospital characteristics and individual-specific characteristics. The model is still a conditional logit, since the only interest is to investigate the change caused on utility by the impact of individual-specific variables on alternative-specific variables.

Finally, the fourth model introduces changes in preferences through time:

$$U_{ijt} = \sum_{t=1}^{T} d_{ij} * \alpha_t + \sum_{t=1}^{T} \sum_{k=1}^{K} q_{jk,t-1} * \beta_{kt} + \sum_{t=1}^{T} \sum_{l=1}^{L} x_{il} * d_{ij} * \delta_{lt} + \\ + \sum_{t=1}^{T} \sum_{k=1}^{K} \sum_{l=1}^{L} q_{j,k,t-1} * x_{il} * \lambda_{klt} + \gamma_j + \varepsilon_{ijt}. \quad (4.9)$$

Each choice is associated with the year in which it was made (interaction terms are inserted for all the considered years minus one, to avoid the dummy variable trap), in order to detect trends and potential changes in patients' perceptions of distance and quality. Each coefficient is now a vector containing T elements: one element is the baseline for year dummy not inserted; the other $T-1$ elements specify the change from year to year. The presence of the interaction terms for time does not interfere with the alternative-specific intercepts: on one side, the baseline level for the independent variables highlights the time-invariant impact of these variables on the choice; on the other side, the intercepts capture the time-invariant impact of unobservable hospital characteristics [101].

All the independent variables have generic coefficients (vectors α_t, β_{kt}, $\delta_l t$, λ_{klt}), which do not depend on the individual or the alternative considered: the effect of each independent variable is constant across all the alternatives and individuals [102]. As an example, the impact of distance on the probability of choice will not depend either on the specific hospital considered, or on the single patient making the choice. At the end of the analysis, the objective characteristics of each alternative, rather than its subjective importance, are used to explain an individual's choice.

4.3 Case study

In order to shed some light on patients' revealed preferences, we applied equation 4.9 to real data, related to more than 15000 patients from the Piedmont Region of Italy that received a treatment for colon cancer surgery between 2004 and 2014. The

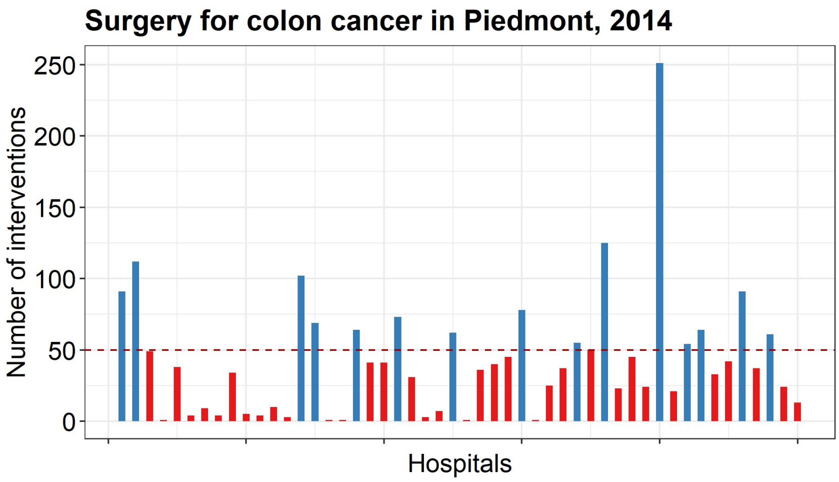


Fig. 4.1 Colon surgeries performed by hospital, 2014

case study raised our interest for multiple reasons, which will be deepened in the following paragraphs.

Colon cancer represents, in Italy, the second most frequent oncological pathology: there were an estimated 37100 cases in 2016, 10% of all the diagnosed cancers [103]. The percentage of patients affected by this disease is increasing: male prevalence raised from 310 (out of 100000 inhabitants) in 2000 up to 450 (female prevalence slightly less, from 225 to 280 out of 100000 inhabitants). For patients diagnosed with colon cancer, a surgical intervention, called resection, is a necessary step in order to remove the cancer. Together with the increase in prevalence, the number of surgeries has increased too: in Piedmont interventions increased from 2067 in 2004 to 2135 in 2014.

Building on the literature about the volume–outcome association, international guidelines recommend that hospitals should provide a minimum of 50 or 70 colon cancer surgeries per year [104]. There have been long-standing concerns that hospitals in Italy are undertaking small volumes of colon cancer surgery, the average number of colon surgeries performed by Italian hospitals in 2015 being 34 [105].

In Piedmont, only the 30% of hospitals (15 out of 50) performing colon surgeries in 2014 had a volumes higher than 50. The number of interventions performed by hospital is shown in Figure 4.1, together with the variability existing among facilities (red bars identify hospitals that perform less than 50 interventions). The situation is better than it used to be, as a result of policies to increase hospital

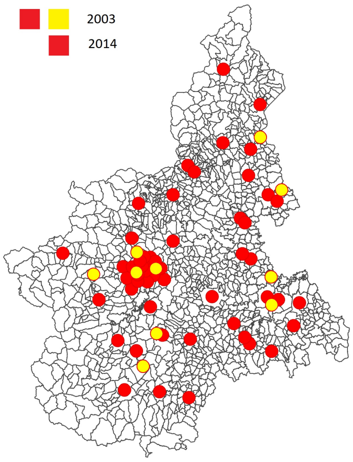


Fig. 4.2 Hospitals location in 2004 (red and yellow dots) and 2014 (red dots)

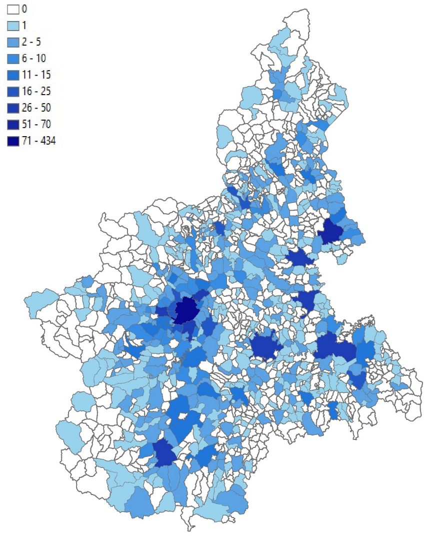


Fig. 4.3 Patients' distribution in Piedmont

specialization. Increased volumes (from an average of 32 in 2004 to 43 in 2014) have been possible through the changes in the regional provision of the service. Figure 4.2 shows differences in hospitals performing colon cancer interventions in 2003 (red and yellow dots) and 2014 (red dots). The comparison with the population density depicted in Figure 4.3 shows that, as expected, hospitals tend to be located in areas of higher population density, even though in some places, notably Turin, more than one hospital provides colon cancer surgery.

Figure 4.4 shows the change in the number of hospitals: the concentration of services has happened gradually, though at a faster pace since 2009. Over the full period, the average number of hospitals in the choice set is 58, starting from 62 in 2004 and arriving to 53 in 2014, a decrease of 15%. The 9 hospitals that stopped performing colon cancer surgery were characterized by an average of 10 interventions per year.

Moreover, at the end of 2015, a regional administrative order [106] officially identified 24 hospitals as hub centres for colon cancer surgery. Figure 4.5 indicates their locations, which correspond to the highly populated centres in the region.

Every city has one hospital providing colon surgery, apart from Turin, where the higher number of hospitals is explained by the larger population (around 20% of the regional population). Even though the administrative order occurred one year after the end of our period of analysis, it is worth of being mentioned since it is the

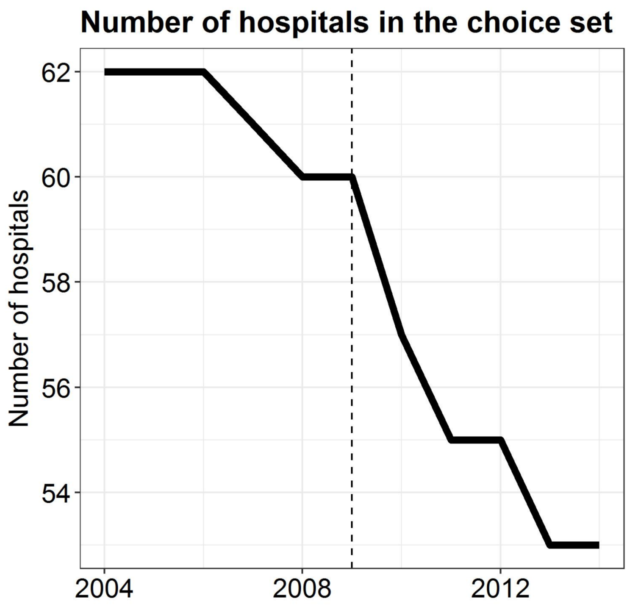


Fig. 4.4 Number of wards operating colon surgery

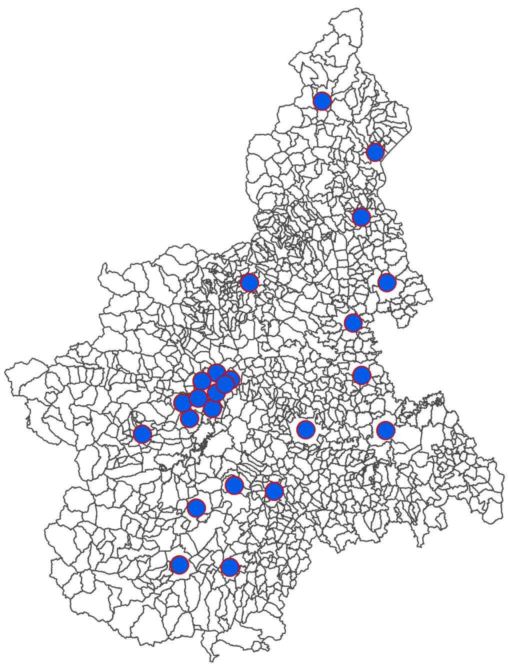


Fig. 4.5 Location of hub hospitals

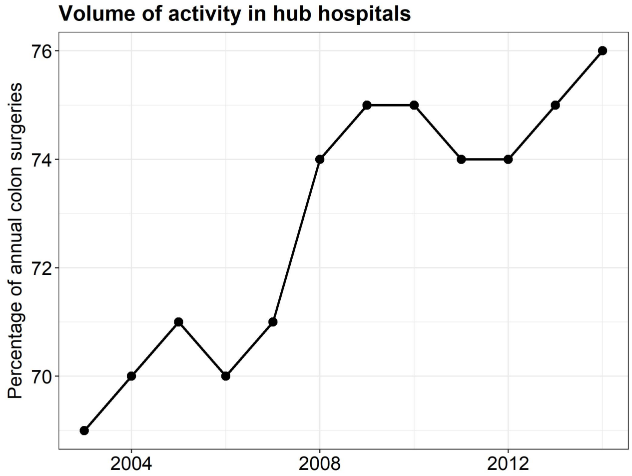


Fig. 4.6 Volume of activity in hub hospitals in Piedmont

final step of a policy towards the concentration of services. The slow transition is demonstrated by Figure 4.6, which documents the constant increase of the percentage of patients being treated by hub hospitals. All in all, the aim of the concentration policy has been to guarantee patients to be treated in more expert facilities.

Even though regional health policy measures are designed to improve patients' clinical pathway, the ultimate decision about where to receive the surgical treatment rests with the patient. In order to clarify when and how patients can exercise their free choice of hospital, we summarize what happens in a patient life before and after the surgical treatment. The path followed by a single patient of our sample is depicted in Figure 4.7.

The first step is the access to the oncological network, which can occur through three channels: (i) if patient age is between 58 and 69 years old, then he/she receives the call from the regional screening programme [107]; (ii) the patient is seen by his/her general practitioner, and the encounter reveals some suspicious symptoms; (iii) the patient, who is not aware of his/her illness, goes to the emergency department because of extreme pain. Once entered the system, the patient undergoes a clinical exam (i.e., endoscopy), which can be performed by the majority of hospitals in Piedmont. If the outcome of the exam detects the presence of colon cancer, the patient receives the diagnosis and a full explanation of the clinical course of the disease, including the need for a surgical intervention. At this point, he/she is invited to choose in which hospital to receive the surgical treatment. The time among the

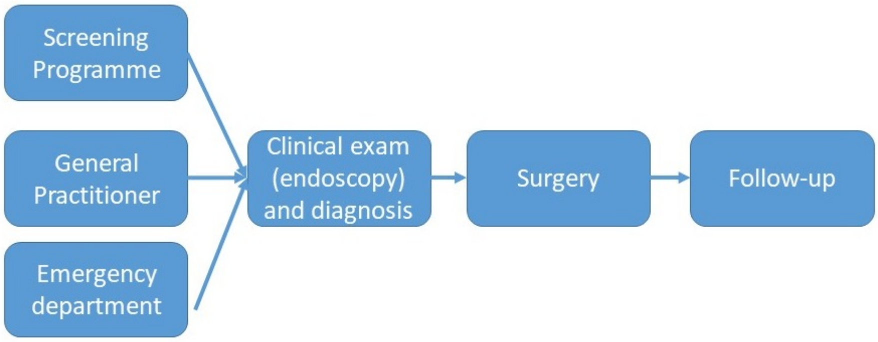


Fig. 4.7 Clinical pathway for colon cancer patients in Piedmont

diagnosis to intervention should not be longer than six months, meaning that patients have time to gather information and develop a personal decision. After the surgery, which requires an average hospital stay of 17 days, the patient receives periodic follow-up by both oncologist and surgeon (who do not need to be the same that have previously treated the patient).

The changes in the regional distribution and organization of hospitals providing colon cancer treatment will have had implications on patient choice. Cessation of activity in some areas reduces the choice set for all the patients and requires some patients to travel further to their nearest hospital. If patients bypass their nearest hospital, we assume this is to receive higher quality care. In order to investigate this assumption, we focus on two measures of hospital quality: (i) waiting times and (ii) mortality. Waiting times are a measure of perceived quality, as patients often prefer to be treated as early as possible. It should be noted, though, that waiting times are not a good proxy for clinical outcomes in colon cancer surgery, since the result of the intervention rarely depends on the waited time (this is true when waiting times are lower than 6 months, which is always the case in Piedmont hospitals). Mortality is a more direct measure of clinical quality, the likelihood of surviving treatment being of primary concern to patients.

4.3.1 Data

Data used in our study are related to two samples, namely a sample of patients and sample of hospitals.

For patients, we use routine administrative patient-level data from the Hospital Discharge Database of Piedmont Region, which were provided by the Epidemiology Unit ASL TO3 of the Piedmont Region. The database contains records of all the episodes of hospitalization. We selected all the Italian residents in Piedmont who were admitted to hospitals located in Piedmont between January 2004 and December 2014, and received an elective intervention for colon cancer surgery (ICD codes for diagnosis 153, 197.5, for intervention 45.7, 45.8, 45.9, 46.03, 46.04, 46.1). We omitted observations on patients younger than 18 years (only 7 cases through the 11 years), because for them the decision may be complicated by parents' influence, whose determinants we do not know. We excluded from the analysis 185 patients (that represent the 0.6% of the sample) who went to private hospitals, because additional elements (e.g., costs) concurred to their choice. The number of observations contained in the data set is 15,663.

We focus on intra-regional mobility, considering only patients who are resident in Piedmont and are treated in Piedmont hospitals. By doing so, we dropped two groups of patients. On one hand, we do not consider patients from other regions that are treated in Piedmont hospitals. On the other, we do not consider patients from Piedmont who went to hospitals in other regions. In this case, we are losing some information, by ignoring the possibility that some patients are willing to travel beyond regional borders. Hence, by considering Piedmont as a closed system, we are underestimating travelled distance, and, as a consequence, the impact of distance on where to receive treatment.

The sample of hospitals was determined through the definition of the choice set, defined on an annual basis. For each year, the set includes all the hospitals in the region that performed colon cancer interventions during the year before. Exceptions are made for (i) hospitals that have been closed: they are excluded from the choice set from the year of closure; (ii) small volume hospitals, with an annual median volume lower than 3 interventions: they are included in the choice set if they performed colon cancer surgery up to three years before, in order to be sure that we can declare the ward as closed. By reducing the choice set in this way, the set includes hospitals about which patients can have some information. However, as a consequence, we exclude from the analysis 163 patients whose chosen hospital is not included in the choice set the year during which they were treated or the year before.

We extract personal information about individual-specific variables ($\sum_{l=1}^{L} x_{ilt}$). More precisely, we look at patient age, sex, residence, comorbidities (measured in terms of Elixhauser comorbid conditions [108]) and rurality. Age is categorized in three levels: baseline (age between 68 and 77), old (age higher than 77), young (age lower than 68). As an indicator of comorbidity, we use a dummy variable, which takes value 1 if the number of Elixhauser comorbidities is greater than 1. In order to assign an indicator of rurality, we associated to each of the municipality of residence the classification made by the Ministry of Economic Development, which categorizes the cities into 6 main groups, ranging from urban (1-3) to rural (4-6) locations. We used a boolean variable with value 0 for the groups 1-3 and value 1 for the group 4-6.

Related to the sample of hospitals, data from the Health Ministry website is used to provide names and addresses for the hospital sites. Further information was extracted by associating information on patients to the hospital that treated them, to construct the alternative-specific variables (d_{ij} and $\sum_{k=1} Kq_{jkt}$). Distance of patient i from hospital j (d_{ij}) was defined as the fastest route by car from the patient's home to each hospital calculated in kilometres [109]. Patients' home was considered to be the centroid of patients' municipality of residence, and the centroid of hospital municipality was considered for its location. An additional and more precise approach (previously done by [110]) was used for patients living in Turin, which is the biggest city of Piedmont. Patients from Turin have the choice set comprising several hospitals within their same city. For these patients, we calculate distance using postal codes, of which there are 33 within the city (elsewhere a single postal code usually applies to each municipality, and 824 (out of 1202) municipalities share their postal codes with another municipality).

Waiting times are measured as the time (in days) elapsed from the day the hospital specialist adds a patient to the waiting list to the day the patient is admitted. Waiting time for a hospital is then calculated as the mean of waiting times undergone by patients treated in that hospital. In-hospital mortality risks are calculated as the percentage of patients that did not survive the period spent in the hospital, out of the total number of patients treated in the same hospital. Indirect risk-adjustment techniques have been used to control for patients' characteristics (such as age, sex, emergency conditions and comorbidities), which affect this measure. Both waiting times and mortality risks are applied to each hospital based on the previous three years performance. This provides a stable measure, particularly for hospitals with

Table 4.1 Descriptive statistics about hospitals' characteristics

	Mean	Median	St. Dev.	5% Quantile	95% Quantile
Waiting time (days)	9.1	7.7	7.1	0.8	21
Mortality risk (% of patients)	4.9	4.3	3.2	0	11.1

small volumes of activity. In particular, for the econometric analysis we further reduced the sample of hospitals to those that performed three year volumes higher than 20 (we defined as a threshold to calculate the mortality risks the three year volume of 20 patients, since the mortality risks for colon cancer are among 3% and 5% in Piedmont during the considered 11 years). Since patients are assumed to judge hospitals based on the previous year information, we dropped 370 patients that chose hospitals for which such information is not available. Hence, the total number of observations is 15,130.

Our database is essentially a pooled cross-section. Since there are no patients appearing in different years, we pool the data in order to maximise sample size and thus increase the precision of estimators. We insert interaction terms that capture variations over time. Interaction terms are the product among independent variables and time dummies (all but one), and they measure the difference between the impact of the independent variable in a specific year and the impact in the baseline (omitted) year.

4.3.2 Descriptive Statistics

Descriptive statistics are reported separately for hospital-specific variables and patient-specific variables.

Hospitals

Table 4.1 illustrates descriptive statistics for the alternative-specific variables, showing that the waiting time for colon cancer surgery averaged 9.1 days over the full period and the mortality risk averaged 4.9%.

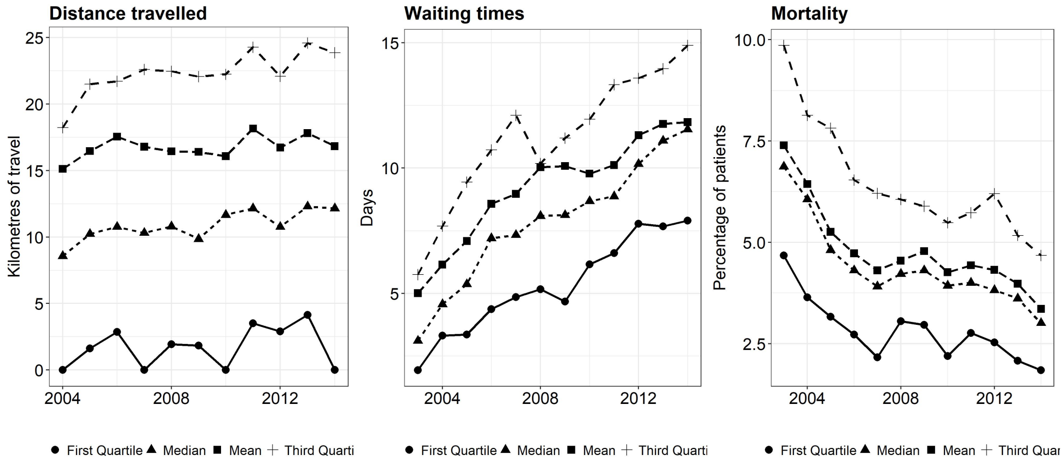


Fig. 4.8 Hospitals' characteristics over time

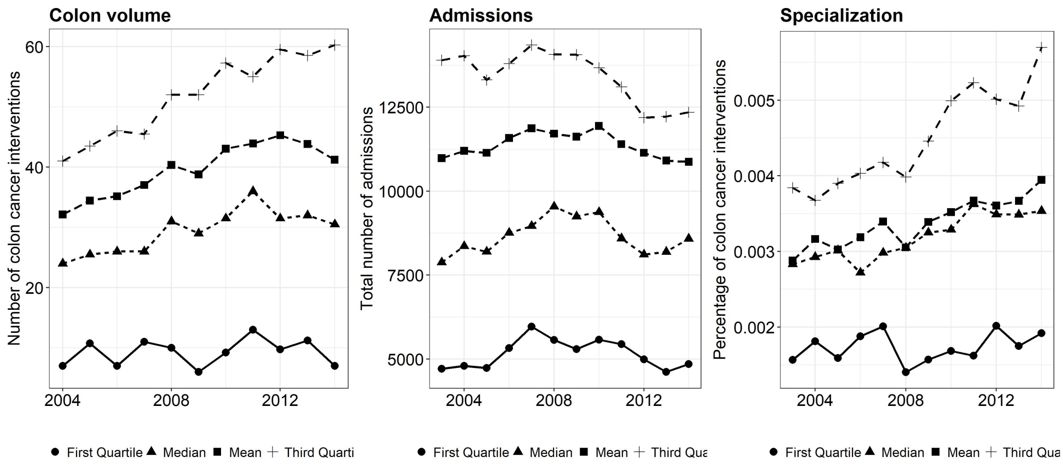


Fig. 4.9 Hospitals' volumes over time

Figure 4.8 shows the trend of hospital characteristics over the 11 years considered, by quartiles of the distribution. Unsurprisingly, given that fewer facilities offer colon cancer surgery, distance travelled has progressively increased. Waiting times have also steadily increased over time, which could be related to the higher concentration of patients in fewer available hospitals. Mortality risk has been constantly decreasing over time, due to (i) improvement in the medical pathway (diagnosis in earlier stage of the cancer, better surgical methods adopted, etc.), (ii) higher specialization of hospitals.

Furthermore, it can be useful to explore the changes that occurred in the volumes performed by hospital, by annually restricting the choice set to the group of hospitals

Table 4.2 Descriptive statistics about patients' characteristics

	Mean	Median	5% Quantile	95% Quantile
Age	70.1	71	52	85
Female	0.46	0	0	1
Number of Elixhauser comorbidities	1.7	1	1	3
Residents in rural municipalities	0.15	0	0	1

that performed some colon surgery. Figure 4.9 illustrates the quartiles between 2004 and 2014 of, respectively, colon surgeries, total admissions and percentage of colon surgeries out of total admissions. As for colon surgeries, even if the first quartile reveals the constant presence of several hospitals performing low volumes, the changes in both mean and median highlight the increased specialization of many other hospitals. The total number of admissions does not show a constant trend, and changes could be due to a number of policies that have been trying to reduce avoidable costs and inappropriate admissions. More interestingly, the third graph reports the increasing trend for the percentage of colon surgeries, which acts as a good proxy of the hospitals specialization. In fact, increase in the index can be due to (i) increase in the number of colon surgeries, which means, increased specialization in the colon intervention; (ii) decrease in the number of total admissions, which means, decreased spread of competences among different specialties. All in all, these trends suggest the positive consequences of the policy that has been adopted in the region.

Patients

Table 4.2 contains descriptive statistics for individual-specific variables. Patients' average age is 71.3. Of them, 46% are female. 15% of patients come from rural areas, and the average Elixhauser comorbidity index is 1.7.

Figure 4.10 shows the characteristics of patients treated over the period. As for age, the number of older people to be operated increased over time, together with a decrease in the number of patients of age between 67 and 77 years old. The trend seems to document the general improvement in the health conditions of the population, which guarantees a longer period of life without clinical surgeries. Male patients have always been more than female ones, even if the percentage has

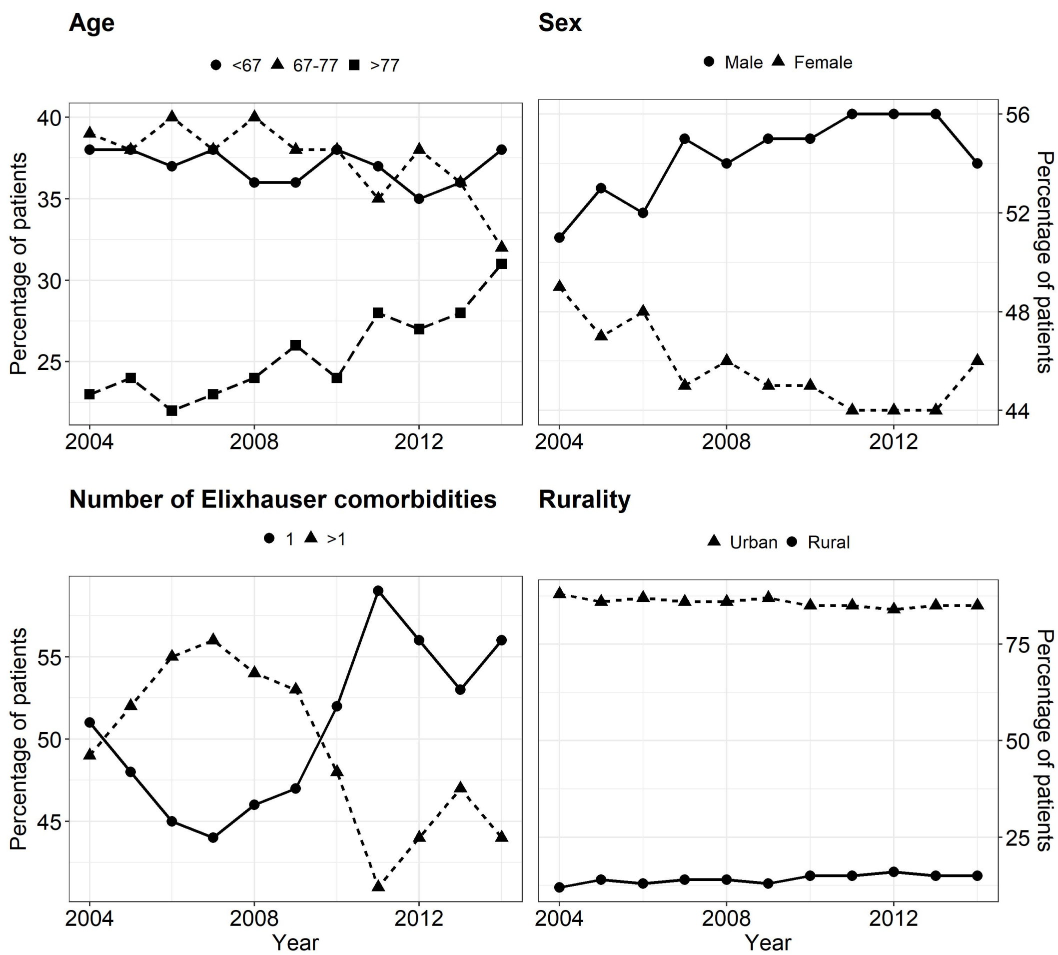


Fig. 4.10 Patients' characteristics over time

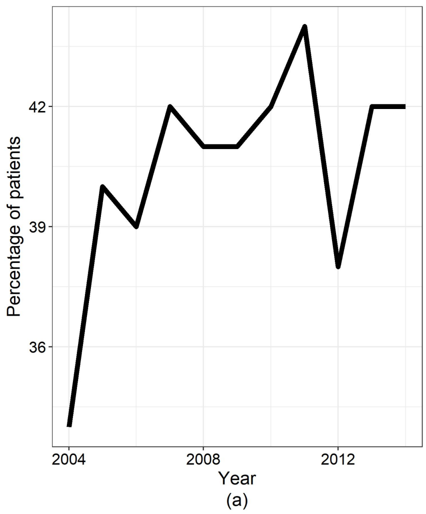


Fig. 4.11 Percentage of bypassers from 2004 to 2014

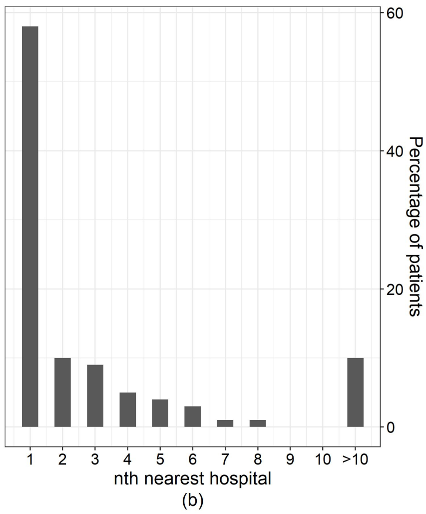


Fig. 4.12 Percentage of patients choosing the nth closest provider, 2014

constantly increased between 2004 and 2011. As for the number of Elixhauser comorbidities, from 2010 the group of patients with only one comorbidity has become bigger than the group of patients with additional comorbidities. As for the classification of municipality, the number of patients from urban areas has had a steeper increase over the period, but the percentages of population that belong to the two categories have remained quite similar over time.

Choice

Patients' revealed preference can be described by examining if and how patients bypassed their nearest hospital. If all the patients went to their closest facility, it can be assumed that only their unwillingness to travel has been considered in making their decision about where to be treated. Figure 4.11 explores the percentage of patients who bypassed their closest hospital, which was on average 41% (similarly to [57]). During the observed period, this percentage increased from 34% in 2004 to 42% in 2014.

In order to understand how far patients travel, Figure 4.12 explores how many patients in 2014 chose their nth closer provider: 58% went to the nearest hospitals. Among the 42% of patients that bypassed their closest hospital, 76% went further than the second closest provider.

Table 4.3 Differences among possible patient choices and actual patient choice

	Mean 2004-2014	Mean 2004	Mean 2014	Median	5% Quantile	95% Quantile
Number of hospitals within 10 km	4	4.5	3.1	1	0	14
Number of hospitals within 30 km	9.8	10	9.1	5	1	21
Number of hospitals within 50 km	16.4	16.8	14.9	18	4	28
Distance to the closest hospital	7.2	6.6	7	5.7	0	21.3
Distance to the chosen hospital	16.1	14.3	16.8	10.9	0	57.7

Differences among possible patient choices and actual patient choice are displayed in Table 4.3. The average (mean) patient has 4 hospitals within a ray of 10 km, 10 within 30 km and 16 within 50 km. The median values, though, are lower, probably because of the sub-set of patients in Turin with a very high number of nearby hospitals. The average distance to the closest hospital is 7.2 km, while that to the chosen hospital is 16 km.

Interestingly, by looking at mean values for the beginning (2004) and the end (2014) of the observed period, we can observe how the characteristics of both the choice set and the choices made by patients changed. On one side, the choice set has been reduced, with fewer hospitals close to patients (e.g., the mean number of hospitals within 10 km decreased from 4.5 in 2004 to 3.1 in 2014). On the other side, patients have been travelling further, on average from 14.3 km in 2004 to 16.8 km in 2014.

By distinguishing among patients that chose the closest hospitals from those who went further, we can observe differences among average characteristics of the chosen hospitals (Figure 4.13).

Over the considered period, obviously, bypassing patients travel further but, surprisingly, they go to hospitals with longer waiting times, although generally lower mortality risks.

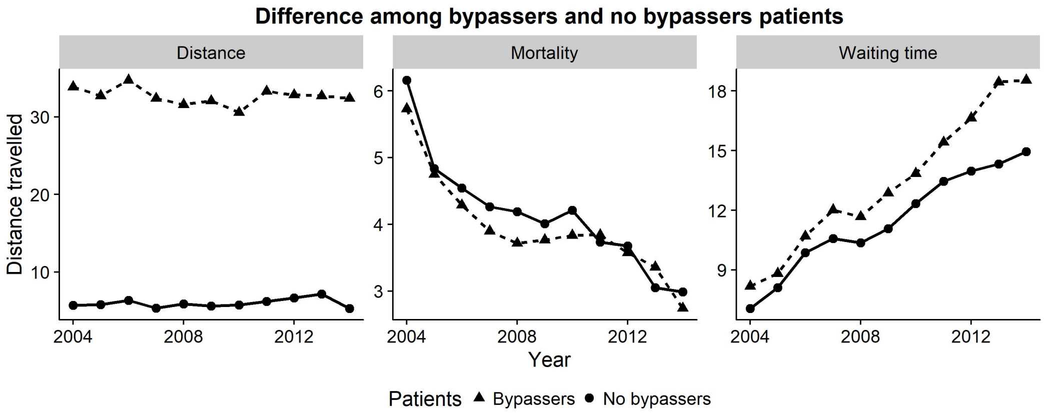


Fig. 4.13 Differences among performance of hospitals chosen by bypassers/no bypassers

Another useful descriptive statistic consists in the analysis of the additional distance travelled by patients, as depending on their personal characteristics. The difference among patients belonging to different categories explains the rational of including individual-specific variables as interaction terms in the conditional logit. In particular, Figure 4.14 shows that the type of patients who appear as willing to travel more are the youngest and rural ones. However, from this graph, we are not able to draw general conclusions, since we are not looking at one variable per time, *ceteris paribus*.

At first sight, data seem to support the hypothesis that patients are making trade-offs in choosing where to have treatment. We examine these trade-offs in Section 4.3.3.

4.3.3 Results

Table 4 reports the coefficients of the regression (details on year interaction terms are shown in Figure 4.15). The coefficient for distance has a negative sign, thus confirming that traveling represents a disutility for patients. Specifically, the ratio of the odds of choosing a hospital that is one km further is 0.9 (=exp (-0.106)). In other words, if hospital B is 1 km further away than hospital A, the odds of the patient going to hospital B will be 0.9 times the odds of the patient going to hospital A.

In order to correctly interpret the coefficients for the quality variables, attention should be paid to the role of alternative-specific intercepts. As explained in [74], by

Table 4.4 Results from the conditional logit

Variable	Est (SE)
Main effects	
Distance	-0.106*** (0.003)
Waiting time	0.021* (0.010)
Mortality Risk	0.016 (0.014)
Interaction with Distance	
Young	0.013*** (0.002)
Old	-0.009*** (0.002)
Female	0.006*** (0.001)
Rural	0.007*** (0.002)
Comorbidity	0.006*** (0.001)
Interaction with Waiting Time	
Young	0.012*** (0.003)
Old	-0.010* (0.004)
Female	0.010** (0.003)
Rural	0.002 (0.005)
Comorbidity	0.001 (0.003)
Interaction with Mortality Risk	
Young	-0.023* (0.009)
Old	-0.001 (0.011)
Female	-0.008 (0.008)
Rural	-0.024* (0.012)
Comorbidity	-0.029*** (0.008)
Alternative-specific intercepts	Y
Year interaction terms	Y
Observations	15,130
Log Likelihood	-26,923
Significance codes: *p<0.05; **p<0.01; ***p<0.005	

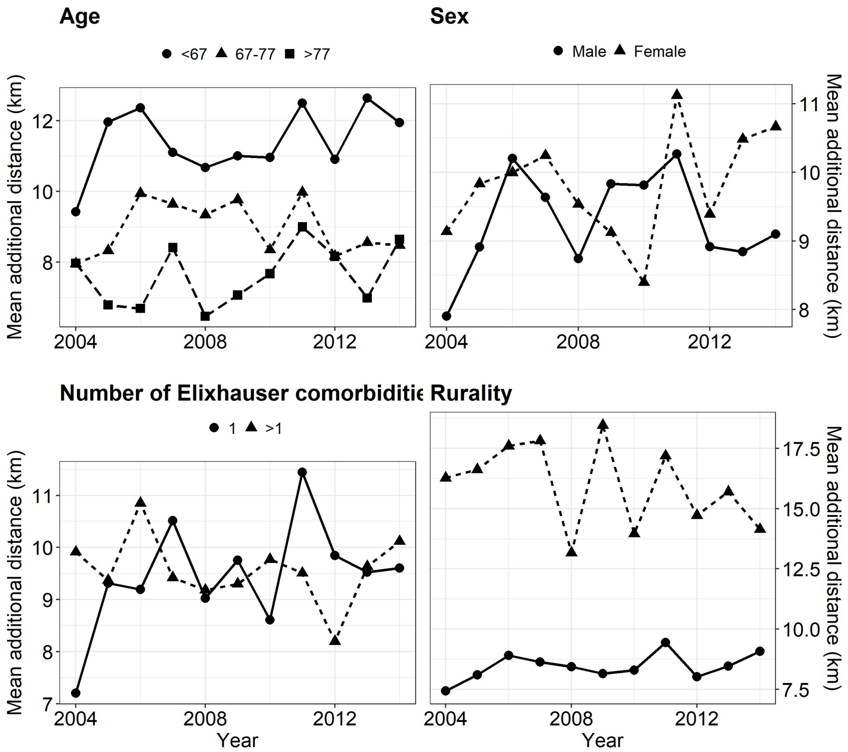


Fig. 4.14 Additional distance travelled for different categories of patients

using intercepts, we split the variation of quality variables from the quality variables themselves.

The baseline coefficient for waiting time is positive and significant, which is contrary to expectations that people would prefer shorter waiting times. However, the median waiting time in the baseline year (2004) was just 3 days. Given such low values, it may be that waiting times above this are a signal of higher quality, representing a small but acceptable delay to access more specialized hospitals.

In contrast, the baseline coefficient for the variation of mortality risks is not significant, suggesting that this information is not informing patient choice.

As for the coefficients of the interactions with individual-specific variables, they provide insights tailored to specific categories of patients. As previously documented in the literature, we find that the disutility caused by distance is higher for older patients [57, 58]). This might be because age is a proxy for frailty or the ability of the patient to travel, with younger, less frail patients being more willing to travel than older patients [60]. A higher propensity to travel is shown also for patients who are

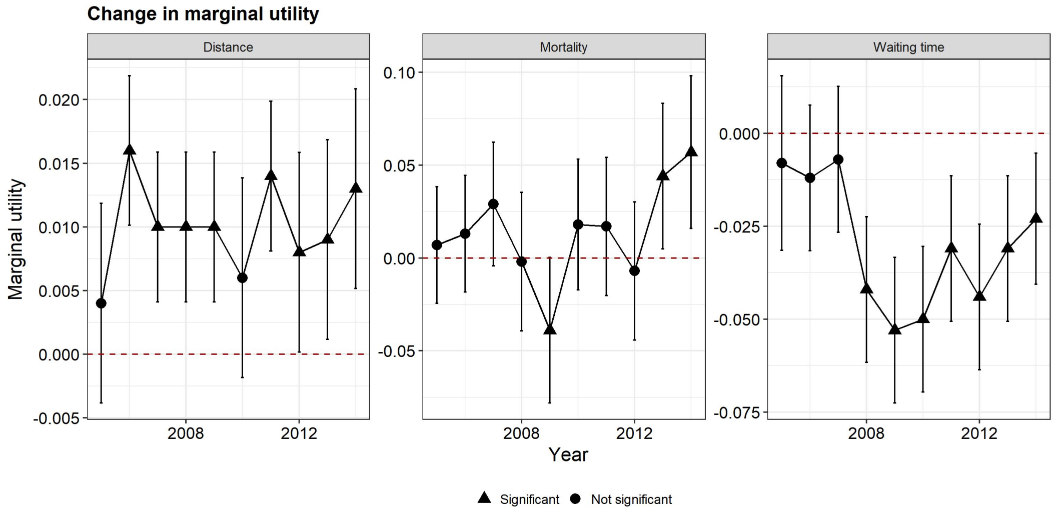


Fig. 4.15 Results on year interaction terms

female, come from rural areas and have more comorbidities. It may be that patients from rural areas are more used to travel in general, and patients with more severe health conditions are more willing to travel in order to be treated.

Interactions among age and waiting times, as well as sex and waiting times, have coefficients with positive signs, signalling that younger and female patients are willing to go to hospitals where they might face longer waiting times. In contrast, older patients tend to choose hospitals with shorter waiting times.

Mortality risks are perceived as a greater source of disutility by patients that are younger, resident in rural areas and, above all, by patients with comorbidities. Instead, patients older than 68, those living in urban areas and those with fewer comorbidities do not consider mortality risks when evaluating their preferences. This could be explained by (i) older patients being less informed, as they face more difficulty in accessing information; (ii) urban patients being less prone to gather information before choosing the hospital, given that it takes less time to evaluate the distance to be travelled; (iii) patients in better health spending less time asking for information about hospitals performance, since they have a lower risk of dying.

Considering the time specifications for each coefficient of the alternative-specific variables, Figure 4.15 shows, compared to the baseline year, the additional change in utility caused by an increase in the alternative-specific variables.

For distance, we observe, apart from 2005 and 2010, significant and positive coefficients. Knowing that the baseline coefficient for distance is negative, these results indicate that patients experienced the highest disutility from distance at the start of the time series. Apart from years 2006 and 2011, disutility associated with distance shows a constantly decreasing trend. This suggests that, contrary to expectations, the policy of concentrating services has not been associated with increased disutility from having to travel further.

For mortality risks, the interaction terms are generally not significant, except in 2009 (negative value), 2013 and 2014 (positive values), where this information does appear to have had a significant and positive effect on patient choice. This unexpected finding merits further investigation.

For waiting times, year interaction terms are significant from 2008 onwards, decidedly so from 2008 to 2010, but progressively less important thereafter. Since all year interaction terms are negative, they reveal the perception of waiting times to be inverted: after 2008, waiting times become a source of disutility for patients. However, as observed for distance, after 2009 there was an upward trend, suggesting recent decreases in disutility associated with waiting times.

We performed robustness checks considering a subset of hospitals: we included only hospitals that performed at least 50 interventions across the whole period considered. As we estimate separate intercepts for each hospital, the risk of including small hospitals is in decreasing the efficiency of the estimates (Gowrisankaran, 1999). Results confirmed what has been obtained from the main analysis.

The results obtained from our analysis shed light on the patient choice process, in particular on how the trade-off among distance and quality has changed during policy roll-out. Overall, the concentration policy did not correspond to an increased disutility perceived by patients. Rather, it has brought about a tendency for patients to travel further and to wait longer. This is somehow reassuring, since whatever reconfiguration can be effective only if patients vote with their feet and choose higher quality hospitals.

4.4 Limitations

The patient model considers patients as single individuals making their own choices. As a consequence, it realistically represents the freedom of choice patients actually have. Nevertheless, by using the patient model, we observe patients to decide (i) independently and (ii) based on information referred to the previous year. Because of these two conditions, the real value of the gained utility could be different from the calculated one. In fact, patients risk to face a disappointing gap among the hospital performances of the previous and the current year. Our statement is built on the assumption that patients' utility is affected by past performance as much as by present performance, since, even though a high volume performed in the past year indicates quality, the volume performed in the current year acts as an additional proof of surgeon's recent experience, and of preparation of the hospital in terms of logistics and organization.

Moreover, because patients choose independently, they lack a territorial overview of the service provided. In fact, patients do not pay attention to the consequences of their choices in terms of quality guaranteed to the whole population. If there is a higher dispersion of demand among facilities, no assurance exists about hospitals performing the threshold number of interventions that allow them to have a sufficient experience. In fact, the institutional figure of the decision maker is the only one devoted to the role of preserving high quality care for the whole population.

Chapter 5

Integrated perspective

In chapters 3 and 4, the perspectives of policy maker and patients have been considered in isolation one from the other. However, in the reality, the two perspectives influence each other.

The objective of this chapter is exactly to take into consideration the existing interaction among patients' choices and decision makers' planning decisions. The first objective is to provide, throughout an integrated approach to the planning problem, a solution that reaches (i) quality in health outcomes and (ii) patients' adherence. The second objective is to investigate the difference among solutions that are identified as optimal by either only one of the actors' perspective, or by considering both the perspectives at the same time.

Figure 5.1 illustrates the processes studied in the previous chapters, together with their interaction. Both of them have inputs (red boxes) and outputs (yellow boxes). As for the decision maker, the input is represented by the actual allocation of interventions among hospitals, which, throughout the process, is transformed

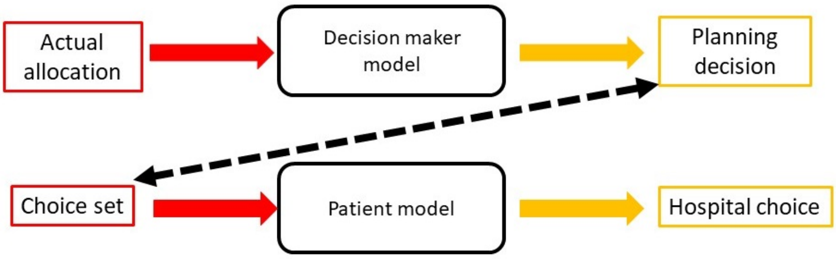


Fig. 5.1 Flow chart describing decision maker and patient models

into a planning decision for a future allocation. As for the patient, the choice set is converted into patients' hospital choice. These two processes are not independent, as the output of the first one is the input of the second one: patients make their hospital choice based on the choice set that has been determined by the policy maker. As a consequence of the interaction among the two perspectives, it is necessary to consider both of them together.

To this purpose, two solution approaches are presented, both aiming to integrate single actors' perspectives. The first approach (that we call *quasi-integrated approach*) is based on a partial integration of the patient perspective into the policy maker model developed in Chapter 3. The second approach (that we call *fully-integrated approach*), instead, tries to achieve a complete integration of the two perspectives by using an algorithmic framework that includes patient choice in the planning decisional process. Since the last decision on planning is made by the decision maker, both the approaches are defined from his/her point of view.

In order to develop the integrated approaches, we refer to the policy maker model and to the patient model formulated in Chapter 3 and Chapter 4. However, since we partially modified them, Section 5.1.1 and 5.1.2 report their updated description. The quasi-integrated and fully-integrated approaches are discussione in Section 5.1.3 and Section 5.1.4, respectively.

Furthermore, we created a framework for a comparative analysis among the approaches, which is introduced in Section 5.1.5. The comparative analysis has the objective of highlighting advantages and disadvantages of each of them, in terms of three categories of performance indicators: mortality, volume and distance.

Eventually, all the models and approaches were applied to the same case study, and compared in Section 5.2. The fully integrated approach emerges as the only one able to overcome the difficulties of the single actors' point of views, by guaranteeing a realistic merge between the planning decision made by the decision maker and the patients' reactions. In fact, it provides a solution that reaches quality in health outcomes and ensures patients to operationally adhere to the planning decision that has been strategically established by the decision maker.

5.1 Problem formulation

As in Chapter 3, we consider a geographic region subdivided in provinces. In each province there is a given number of hospitals and of patients. However, we now focus on a single surgical intervention, and we consider only the facilities performing that intervention. As in Chapter 3, we assume to have access to information about (i) hospitals past interventions, i.e., we are aware of the past performance in terms of clinical outcomes; (ii) the total number of patients who require to be operated in a given year, as well as their provenance and personal characteristics.

In the following sections, the different perspectives will be considered. Section 5.1.1 and 5.1.2 will shortly recap the points of view of decision maker and patient, starting from the models presented in Chapter 3 and 4, and highlighting the required changes.

5.1.1 Policy maker model

The first perspective considered is the one of the policy maker, which focuses on the decision of how to allocate surgery interventions among the hospitals operating in a same geographical area. The model originates from the one presented in Chapter 3, with some changes that will be explained in detail.

As discussed in Section 3.1, when considering the allocation problem, attention must be paid to the factors that can bound (from above or from below) the number of interventions allocated to a hospital. In fact, in order not to overload the structure, or to cause burn-out to the personnel, each hospital ward is considered to have a maximum capacity (in terms of performable interventions) that cannot be exceeded. However, at the same time, international guidelines advocate for a minimum number of interventions that need to be performed in a hospital in order to guarantee sufficient staff experience, and, as a consequence, adequate quality levels. These two opposite requirements specify the upper bound and lower bound, respectively, of the number of operations that can be allocated to a hospital. Compared to the previous formulation of the policy maker model, here we do not only consider the hospital capacity, rather we add the relevance of the internationally suggested thresholds of volume.

Moreover, we make use of the information about hospital past performance, by attributing to them the performance coefficients presented in Section 3.4.4. In this way, we allow hospital outcomes to vary, depending on surgeons' skills revealed in the previous year.

A new constraint is added in order to regulate the provenance of the patients treated in a hospital: we allow hospitals of a given province to treat also patients coming from provinces other than theirs, and, at the same time, we set a threshold on the interventions that are performed to $foreigner$. This constraint enables us, for the sake of simplicity, to focus on the extreme case where each province takes care only of its own patients.

About the demand, as in Section 3.1, we assume that all the patients that require an intervention must receive it, and we set a fixed regional funding as upper bound for the costs of the interventions that can be borne. Given these constraints and assumptions, the policy maker objective is to allocate the volumes of interventions among health facilities in order to improve healthcare quality, i.e., to reduce total mortality.

The used notation is reported in Table 5.1. Variables and parameters that are introduced for the first time will be explicitly mentioned. Using the notation of Table 5.1, the problem can be formulated as follows:

$$\min \quad z_{policy} = \sum_{n=1}^{N} \sum_{j=1}^{J_{nt}} x_{j_n,t} * m(x_{j_n,t}) + \sum_{n=1}^{N} \sum_{j=1}^{J_{nt}} x_{j_n,t} * v_{j_n,t} * dev(x_{j_n,t}) \tag{5.1}$$

$$\text{s.t.} \quad f_{j_n,t} * T \leq x_{j_n,t} \leq f_{j_n,t} * cap_{j_n,t} \qquad \forall n, j_n, t \tag{5.2}$$

$$x_{j_n,t-1} \geq T \qquad \forall n, j_n, t \tag{5.3}$$

$$I_{nt} - \delta * I_{nt} * \sum_{n'=1, n' \neq n}^{N-1} 1 \leq \sum_{j=1}^{J_{nt}} x_{j_n,t} \leq I_{nt} + \delta * I_{nt} * \sum_{n'=1, n' \neq n}^{N-1} 1 \qquad \forall n, t \tag{5.4}$$

$$\sum_{n=1}^{N} \sum_{j=1}^{J_{nt}} x_{j_n,t} = \sum_{n=1}^{N} I_{nt} \qquad \forall t \tag{5.5}$$

$$\sum_{n=1}^{N} \sum_{j=1}^{J_{nt}} x_{j_n,t} * c \leq b \qquad \forall t \tag{5.6}$$

$$x_{j_nt} \in Z^+, f_{j_n} \in \{0,1\} \qquad \forall n = 1...N, j = 1...J_n, t$$

Table 5.1 Notation used for the mathematical formulation of the policy maker model

Indexes	
Name	**Definition**
j	Hospital
n	Province
t	Year
Decision variables	
Name	**Definition**
$f_{j_n,t}$	Boolean variables, which assume value 0 if hospital j of province n is closed in year t, 1 if it is open
$x_{j_n,t}$	Integer variables, which indicate the total volume performed by hospital j of province n in year t
Parameters	
Name	**Definition**
I_{nt}	Total number of patients in province n in year t
J_{nt}	Total number of hospitals in province n in year t
N	Total number of provinces
$cap_{j_n,t}$	Capacity of hospital j of province n in year t
δ	*Hospitality threshold*: identifies the threshold of patients (in percentage) from other provinces that can be treated in the hospitals of a province
c	Marginal cost for each surgical intervention
b	Available regional funding
$m(x_{j_n,t})$	Mortality risk of patients being treated in hospital j of province n in year t
$v_{j_n,t}$	Performance coefficient of hospital j of province n in year t
T	Threshold of interventions suggested by international guidelines
$dev_{x_{j_n,t}}$	Average deviation from the mortality curve corresponding to the volume $x_{j_n,t}$

It should be noticed that, compared to Chapter 3, we inserted an additional index, t, which specifies the year in which the allocation is made. Even though we focus on the decisional process occurring in a unique year t, the time specification is useful for (i) one additional constraint (5.3) inserted for the policy maker model, which will be explained in the following paragraphs; (ii) the integration that will be made in Section 5.1.3. Moreover, compared to Chapter 3, we label hospitals with the index, j_n, in order to specify the province n where each hospital j is located.

The objective function (5.1) refers to the equation (3.15), and expresses the minimization of the total mortality (related to the considered intervention) on the territory, taking into account the hospital performance coefficients. In fact, the first term is the volume of activity multiplied for the mortality risk of the hospital (which in turn depends on its performed volume, due to the volume–outcome association), while the second term has the role of increasing/decreasing the mortality depending on the performance coefficient of each hospital.

Differently from Chapter 3, here we insert boolean variables, $f_{j_n,t}$, which describe the opening or closure of hospitals. For each hospital, the volume operated either can be zero if the hospital is closed (i.e., if $f_{j_n,t} = 0$), or, if the hospital is open, it has to be larger than the T threshold, but smaller than its capacity (equation 5.2). Notice that all the hospitals that have a capacity lower than T (i.e., the threshold of interventions suggested by international guidelines) are not considered in the model since they would lead to an infeasible solution. The same happens for all the hospitals that did not perform the T threshold in the previous year (equation 5.3). This is a further action (compared to Chapter 3) that is undertaken to concentrate the regional demand into wards that have a cumulated a sufficient level of experience.

The new set of constraints added to regulate the provenance of patients treated in a hospital are expressed in the equation (5.4). They consider the demand satisfaction from the perspective of the province, by looking at the total volume performed by the hospitals that are in the same province. The inequalities allow hospitals to treat, beyond their provincial population, a *hospitality threshold* of patients coming from different provinces. Vice versa, if patients from province A are accepted in the hospitals of other provinces, the hospitals located in province A are allowed to treat less cases than their provincial population (up to the extreme case where the other provinces treat all the patients from province A). In the equation (5.4), the left part represents the outflow of patients (namely, the decrease in provincial demand due to patients treated in the hospitals of other provinces), while the right part considers the inflow of patients (i.e., patients from other provinces being treated in the hospital). If the hospitality threshold assumes null value, the hospitals in a province have to treat all the patients living in that province and requiring an intervention. Hence, in principle, the treatment of the patients from a province represents a duty for the hospitals of the same province.

Eventually, as in the policy maker model, regardless of how the demand is scattered among facilities, it has to be fully satisfied, i.e., there should not be patients needing the treatment who are not treated (i.e., under-treatment), or patients that receive the treatment without needing it (i.e., over-treatment) (equation (5.5)); and the available budget has not to be exceeded (equation (5.6)).

We call the solution of the policy maker model, in terms of volumes assigned to hospitals, *strategical solution*.

5.1.2 Patient model

In this section, we consider the perspective of a single patient, who chooses what is the best hospital to be treated in. The methodology used to this purpose is the same presented in Chapter 4, i.e., it is a choice model, based on the conditional logit random utility model. However, while in Chapter 4 we were interested in modelling the decisional process that had been undertaken to reach patient choice, in this section we aim to predict patient choice. Given a configuration of hospitals, each characterized by a location and a quality attribute, we want to observe how patients would distribute among them.

The choice of a patient for a specific hospital has been considered as depending both (i) on the value perceived from receiving care in that hospital, which means, on the expected utility gained, and (ii) on the comparison among expected utilities that would be gained if seeking care from different available hospitals. We assume that patients behave so as to maximize their expected utility.

Patient's utility is expressed as the sum of two components, referring to distance and quality, respectively. Among the different quality measures that have been mentioned in Section 2.2, we use the volume of activity: the higher number of interventions the hospital performs, the stronger the signal for specialization, experience and thus quality. Hence, in the following, the two terms (quality and volume) will be used as synonymous. Furthermore, patients' personal characteristics, such as age, sex, provenance and health conditions, affect the perception of distance and quality.

As in the policy maker model, we consider a given region with its provinces, each of which accounting for a given number of patients, and containing a given number of hospitals. Since we take the patient perspective, all the hospitals in the region are considered part of patients' choice set, meaning that they are assumed available to be chosen, despite the province in which they are located and the province in which the patient lives.

Several identical models, each one referred to a different patient living in a specific province, are needed to represent the decisional process of every single patient. Each model analyses the propensity for a patient to choose a specific hospital, depending on both the preferences of the patient, and the characteristics of the hospital, as observed by the patient during the year before his/her choice. Eventually, the distribution of volumes among hospitals is determined by aggregating patients' behaviour.

Since the parameters used in this model are the same as the previous chapter, they are shortly presented in Table 5.2. It should be noticed that, compared to the decision maker model, here the volume by facility, $x_{j_n,t}$, is a parameter instead of a variable.

The patient model is represented by the following equations:

$$p_{i_n,j_n,t} = \frac{U_{i_n,j_n,t}}{\sum_{n=1}^{N}\sum_{j=1}^{J_{nt}} U_{i_n,j_n,t}} \forall n, j_n, t \tag{5.7}$$

$$U_{i_n,j_n,t} = \left(\beta_{d1} + \sum_{k=1}^{K} \beta_{kd2} * g_{ki_n}\right) * d_{i_n,j_n} + \\ + \left(\beta_{q1} + \sum_{k=1}^{K} \beta_{kq2} * g_{ki_n}\right) * x_{j_n,t-1} \quad \forall n, j_{nt}, t \tag{5.8}$$

$$\sum_{n=1}^{N}\sum_{j=1}^{J_{nt}} p_{i_n,j_n,t} = 1 \tag{5.9}$$

$$p_{i_n,j_n,t} \in \{0,1\} \qquad \forall n = 1...N, i = 1...I_{nt}, t$$

For each model of patient i from province n in year t, the decision variables are $p_{i_n,j_n,t}$, which represent the probability of patient i of province n to choose hospital

Table 5.2 Notation used for the mathematical formulation of the patient model

Indexes	
Name	**Definition**
i	Patient
j	Hospital
n	Province
t	Year
Decision variables	
Name	**Definition**
$p_{i_n,j_n,t}$	Probability of patient i of province n of choosing hospital j of province n in year t
$U_{i_n,j_n,t}$	Utility gained by patient i of province n by choosing hospital j of province n in year t
Parameters	
Name	**Definition**
d_{i_n,j_n}	Distance separating patient i of province n from hospital j of province n
$x_{j_n,t}$	Volume that is allocated to hospital j of province n in year t (quality measure)
β_{q1}	Increase in utility due to an additional intervention previously performed in the chosen hospital
β_{kq2}	Change in β_{q1} that is due to the personal characteristic k of a patient
β_{d1}	Increase in utility caused by one additional kilometre travelled to reach the chosen hospital
β_{kd2}	Change in β_{d1} that is due to the personal characteristic k of a patient
g_{ki_n}	Information about the k personal characteristic of patient i from province n (i.e., age, sex, rurality and comorbidity)

j of province n in year t. This probability expresses the patient propensity for a hospital.

Equation 5.7 expresses patient's behaviour, which is entirely centred on the perceived utility. In fact, the probability of a patient to choose a specific hospital will increase if the utility perceived by his/her choice increases and/or the sum of the utilities gained by choosing all the hospitals decreases.

Equation (5.8) shows how the patient utility is computed, i.e., summing up the two main components related to distance and quality. As already done in Chapter 4, patients have the possibility of gathering only past information about hospital quality. The magnitude of the impact of distance and quality on the utility is given by the β coefficients, and by their interactions with personal characteristics. We assume that patients do not face any strict constraint as for maximum distance travelled and minimum quality required. Rather, the preference function already includes all the aspects of their decisional process.

Equation (5.9) bounds the sum of the probabilities of going to any hospital to be 1. In this way, we represent the link existing among the chosen hospital and all the available ones: the preference for each hospital affects the final patient choice.

As a result, we obtain the probability of every patient of choosing every hospital in every year. In order to predict the yearly hospital volume, the following formula can be used:

$$x_{j_n,t} = \sum_{n=1}^{N} \sum_{i=1}^{I_{nt}} p_{i_n,j_n,t} \tag{5.10}$$
$$\forall \quad n = 1...N, j = 1...J_n, t$$

Equation (5.10) shows that, for every hospital of every province in every year, the total predicted number of patients choosing it will be the sum of the probabilities of each patient choosing it. This equation can be read also as the multiplication of the total number of patients by the average probability of all the patients of choosing that hospital. We call the solution of the patient model, in terms of volumes predicted to be assigned to hospitals, *operational solution*.

5.1.3 Quasi-integrated approach

The first integrated approach is based on the policy maker model. However, differently from the policy maker model, where only the total volumes (i.e., the number of interventions) performed by hospitals were considered, here the same total volume is split into the sum of single actions made by patients. In other words, the decision maker is not any more looking for the optimal solution only at an aggregated level, rather he/she is deciding which are the best individual choices that patients have to do. Patients' answer, in terms of their adherence to the plan established by the decision maker, will be discussed at the end of the paragraph.

In addition to the disaggregated total volumes, the first integrated approach also includes the maximum distance patients are willing to travel, which is calculated through the patient model. Since distance is a source of disutility, every patient will choose his/her closest hospital, unless the quality provided by a a facility compensates the higher distance required to reach it. Each patient has a propensity to travel, which can be expressed as the amount of additional kilometres patients are ready to cover in order to reach a hospital of additional quality. In terms of the model described in Chapter 4, patients' willingness to travel compares the utility that patients gain with quality to the utility that patients loose with distance.

In the quasi-integrated approach, the decision maker is in charge of guaranteeing to all the patients to go in hospitals that are either the closest, or at a distance that is offset by the higher quality provided, given their personal preference to distance and quality.

To model this problem, new variables and parameters are required with respect to the policy maker model introduced in Section 5.1.1. In fact, in this case, the decision variables are, beyond the boolean variables $f_{j_n,t}$ concerning the opening/closure of hospital j of province n in year t, the $y_{i_n,j_n,t}$ binary variables, which indicate if patient i of province n goes in hospital j of province n in year t.

The parameters refer both to hospitals and patients and are basically the ones used in the previous sections, with the following exceptions: (i) the volume by facility, $x_{j_n,t}$, is not any more a variable; (ii) the two parameters min_{i_n,j_n} and wtt_{i_n} are introduced, representing the closest hospitals for patients and the propensity to travel, respectively.

The used notation is reported in Table 5.3.

Table 5.3 Notation used for the quasi-integrated approach

Indexes	
Name	**Definition**
i	Patient
j	Hospital
n	Province
t	Year
Decision variables	
Name	**Definition**
$f_{j_n,t}$	Boolean variables, assuming value 0 if hospital j of province n is closed in year t, 1 if it is open
$y_{i_n,j_n,t}$	Boolean variables, indicating if patient i of province n is treated in hospital j of province n in year t
Parameters	
Name	**Definition**
I_{nt}	Total number of patients in province n in year t
J_{nt}	Total number of hospitals in province n in year t
N	Total number of provinces
$x_{j_n,t}$	Volume that is allocated to hospital j of province n in year t
$cap_{j_n,t}$	Capacity of hospital j of province n in year t
δ	*Hospitality threshold*: identifies the threshold of patients (in percentage) from other provinces that can be treated in the hospitals of a province
c	Marginal cost for each surgical intervention
b	Available regional funding
$m(x_{j_n,t})$	Mortality risk of patients being treated in hospital j of province n in year t
$v_{j_n,t}$	Performance coefficient of hospital j of province n in year t
T	Threshold of interventions suggested by international guidelines
$dev_{x_{j_n,t}}$	Average deviation from the mortality curve corresponding to the volume $x_{j_n,t}$
d_{i_n,j_n}	Distance separating patient i of province n from hospital j of province n
min_{i_n,j_n}	Boolean variables, which assume value 1 if hospital j of province n is the closest to patient i of province n, 0 if not
β_{q1}	Increase in utility due to an additional intervention previously performed in the chosen hospital
β_{kq2}	Change in β_{q1} that is due to the personal characteristic k of a patient
β_{d1}	Increase in utility caused by one additional kilometre travelled to reach the chosen hospital
β_{kd2}	Change in β_{d1} that is due to the personal characteristic k of a patient
g_{ki_n}	Information about the k personal characteristic of patient i from province n (i.e., age, sex, rurality and comorbidity)
wtt_{i_n}	Propensity to travel of patient i from province n
α_{wtt}	Coefficient allowed for the relaxation of the constraint on the willingness to travel

$$\min \quad z_{policy} = \sum_{n=1}^{N}\sum_{j=1}^{J_{nt}}\sum_{i=1}^{I_{nt}} y_{i_n j_n,t} * m\left(\sum_{n=1}^{N}\sum_{i=1}^{I_{nt}} y_{i_n j_n,t}\right) +$$
$$+\sum_{n=1}^{N}\sum_{j=1}^{J_{nt}}\sum_{i=1}^{I_{nt}} y_{i_n j_n,t} * v_{j_n,t} * dev\left(\sum_{n=1}^{N}\sum_{i=1}^{I_{nt}} y_{i_n j_n,t}\right) \tag{5.11}$$

$$\text{s.t.} \quad f_{j_n,t} * T \leq x_{j_n,t} \leq f_{j_n,t} * cap_{j_n,t} \qquad \forall n, j_n, t \tag{5.12}$$

$$I_{nt} - \delta * I_{nt} * \sum_{n'=1,n'\neq n}^{N-1} 1 \leq \sum_{j=1}^{J_{nt}} x_{j_n,t} \leq I_{nt} + \delta * I_{nt} * \sum_{n'=1,n'\neq n}^{N-1} 1 \qquad \forall n,t \tag{5.13}$$

$$\sum_{n=1}^{N}\sum_{j=1}^{J_{nt}} x_{j_n,t} = \sum_{n=1}^{N} I_{nt} \qquad \forall t \tag{5.14}$$

$$\sum_{n=1}^{N}\sum_{j=1}^{J_{nt}} x_{j_n,t} * c \leq b \qquad \forall t \tag{5.15}$$

$$x_{j_n,t} = \sum_{i=1}^{I_{nt}} y_{i_n j_n,t} \qquad \forall n, j_n, t \tag{5.16}$$

$$y_{i_n j_n,t} * \left(d_{i_n j_n} - \sum_{n=1}^{N}\sum_{j=1}^{J_{nt}} d_{i_n j_n} * min_{i_n j_n}\right) \leq$$
$$\leq \alpha_{wtt} * wtt_{i_n} * \left(x_{j_n,t} - \sum_{n=1}^{N}\sum_{j=1}^{J_{nt}} x_{j_n,t} * min_{i_n j_n}\right) \forall t \tag{5.17}$$

$$wtt_{i_n} = -\frac{\beta_{d1} + \sum_{k=1}^{K} \beta_{kd2} * g_{ki_n}}{\beta_{q1} + \sum_{k=1}^{K} \beta_{kq2} * g_{ki_n}} \tag{5.18}$$

$$x_{j_n,t} \in Z^+, f_{j_n,t} \in \{0,1\} \qquad \forall n = 1...N, j = 1...J_{nt}, t$$

The integrated model has the same objective function (5.20) and constraints (5.14, 5.15, 5.16) as the policy maker model presented in 5.1.1. However, the volume of activity of each hospital j of province n in year t, $x_{j_n,t}$, is now considered as the sum of the number of patients choosing to be operated in it (equation (5.16)), i.e., the decision maker now takes into consideration single actions performed by patients who travel to a specific facility.

Constraints (5.17) represent the maximum distance patients are willing to travel. Since, as shown in Section 4.2, the utility of patients is linearly expressed as the sum of two components, distance and quality, we can calculate for each patient (through the specification of his/her K personal characteristics) the ratio of marginal utility gained by the distance out of marginal utility gained by the quality (equation (5.18)). This ratio can be interpreted as the willingness to travel, i.e., the amount of additional kilometres patients are ready to cover in order to reach a hospital of additional quality. Equations (5.17) translate patients' propensity to travel into a constraint. Specifically, compared to the closest hospital, the additional amount of kilometres patients have to cover to reach the chosen facility cannot exceed the increase in quality provided in the chosen hospitals (in our study, the increase in the number of operations performed), multiplied by the factor wtt_{i_n}. In other words, since distance is a source of disutility for patients, there must be an increase in quality that offsets the travelled kilometres. Patients will choose the closest hospital, unless the quality provided by another facility compensates the higher distance required. By using the parameter α_{wtt}, we accept the possibility of partially relaxing this constraint, which could be useful for a more immediate implementation of the model, as it will be clear in the following.

Considering the volume as the union of single patients' choices, the decision maker is implicitly also deciding which patients are going where, beyond the total volume allocated to hospitals. Hence, to further stress the decision maker attention towards the population wellbeing, we decided to allocate patients to hospital in a non-random way. In particular, we assume that the decision about the allocation is driven by the *closest hospital* criteria, i.e., the decision maker aims at each patient going to the closest hospital so as to minimize distance travelled and, hence, the disutility perceived by patients. However, whenever the closest hospitals are not able to fulfil the entire demand from their catchment area, some patients are forced to bypass their closest hospital and to move to the next one.

To solve the problem, we developed a solution algorithm that selects the patients forced to bypass based on their willingness to travel, by choosing the ones with the highest values. In this way, the probability of the constraints (5.17) to hold is increased, and the patients who are forced to travel are the ones who perceive the lower disutility from mobility. For the same reason, when multiple hospitals are present in the same municipality and their total capacity is higher than the population in their catchment area, the algorithm leaves free space in the hospitals to which higher volumes have been assigned. In fact, the patients from other provinces who are forced to bypass, will be more likely to accept to travel to reach the hospitals

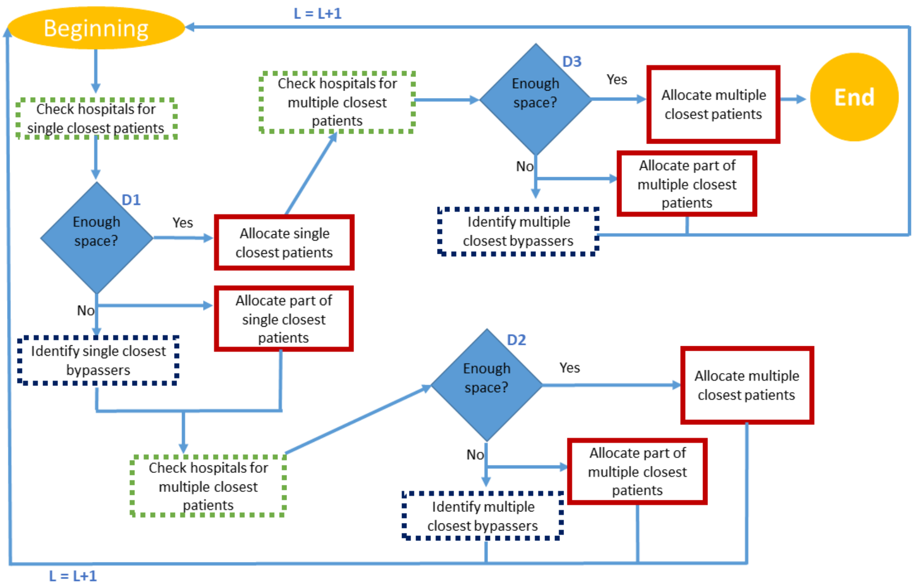


Fig. 5.2 Solution algorithm for the quasi-integrated approach

with higher quality levels. All in all, since a less distance travelled and a higher quality are associated to patients' utility, the developed algorithm tries to increase the utility that the decision maker is guaranteeing to the covered population.

The flowchart of the solution algorithm is reported in Figure 5.2. It is an iterative algorithm. In each iteration L, the L-th nearest hospital to the still not assigned patients are considered. At the beginning of each iteration, a distinction is made between the patients who have only one closest hospitals (that will be called *single closest patients*) and the patients who have multiple hospitals at the same, shortest, distance (that will be called *multiple closest patients*). Single closest patients are considered as first, since they have fewer possibilities to travel as less as possible. The blue diamond $D1$ is the capacity check, i.e., to control if hospitals that are the closest for single closest patients have enough space to treat all of them.

If the check is negative, which corresponds to full hospitals that cannot treat all the first closest patients, a groups of patients need to be chosen to become *bypassers*, i.e., to travel farther to another hospital. To this purpose, the closest hospital criteria is applied: for each overloaded hospital, the group of closest patients with the highest willingness to travel is chosen. The size of the group is equal to the amount of interventions that cause the overload. In this way, we are choosing as bypassers those patients that are more prone than others to travel.

Together with the identification of the bypassers, we have also defined which patients are to be directed to the L-th closest hospital. Then, the same capacity check is repeated for the hospitals that are closest to multiple closest hospitals (diamond $D2$):

- If the answer is negative, multiple closest bypassers are identified using the same criteria used for single closest patients. Then, all the patients that are bypassers are grouped, the non-bypassers are allocated to their L-th closest hospital, and a new iteration is started.
- If the answer is positive, the multiple closest patients are allocated to one of their closest hospital. In order to decide to which ones, if the capacity of all the subgroups of hospitals that are the closest to multiple closest patients is higher than the number of the multiple closest patients, they are ranked. The ranking is based on the number of interventions that have been strategically assigned to them by the decision maker, i.e., the first hospital will be the one with the highest volume planned by the decision maker. The aim is to leave free space to the hospitals that have higher rankings, so multiple closest patients are allocated to the low-rank hospitals. In this way, the single closest patients that have been previously chosen as bypassers (and, hence, that will be redirected to one of these hospitals with free capacity) will gain higher utility caused by higher quality, which compensates the higher disutility caused by the higher distance. After this allocation, a new iteration is started.

On the contrary, if the check D1 is positive, which corresponds to hospitals able to treat all the first single closest patients, all the single closest patients are allocated. Then, the closest hospitals for the multiple closest patients are identified and their capacity check (diamond $D3$) is performed.

- If the capacity is not sufficient, a group of multiple closest bypassers is identified in the same way as already done for single closest patients. Then, a new iteration starts.
- If the capacity is sufficient, all the multiple closest patients are allocated to their closest hospitals, and, since all the patients have been allocated, the algorithm is ended.

It should be noticed that, in order to consider the assignment of specific patients to specific hospitals, we preferred to develop a solution algorithm, rather than to add a constraint to the integrated model. In fact, we have reckoned the risk of the *closest hospital* criteria to increase the complexity of the model, since it requires an iterative approach that gradually assigns patients to hospitals with available capacity. Our methodological choice has been then to extract the aggregated volumes assigned to hospitals from the integrated model, and to split them into single patient assignments through the described solution approach.

Regardless the algorithm used to solve the model, it can be noticed that, by adding the constraints (5.15, 5.16, 5.17), the decision maker takes charge of guaranteeing patients to travel no more than they are willing to. Nevertheless, by using patients' choices as decision maker's decision variables, we are still assuming that patients will go to the hospital that has been thought for them. Actually, talking of choices, it would imply to leave space to patients' own decisional process, which is not what happens with the quasi–integrated approach (hence, the name *quasi*). This case will be treated in the following section.

5.1.4 Fully-integrated approach

The fully-integrated approach aims at a realistic merge among policy maker and patient perspectives. Since the last decision on the planning choice is made by the policy maker, the fully-integrated approach is defined using his/her point of view, exactly as the quasi-integrated approach. However, in this case, the policy maker does not impose his/her choice on patients, neither on an aggregate level (as in the decision maker model), nor on an individual level (as in the quasi-integrated approach). The aim is to merge the objective of low mortality with a positive patients' adherence, so that the fully-integrated approach includes patients' answer. In other words, the fully-integrated approach represents the point of view of a policy maker who, in order to decide how to allocate the interventions to the hospitals, takes into account also how patients will behave.

In particular, the approach is based on the prediction of patients' behaviours, which becomes included in the decisional process of the policy maker. By doing so, we remove the assumption of patients' total adherence to the decision maker plans.

The main idea is to control *ex-ante* how patients would react to different choice sets, in terms of the probability of choosing a particular hospital within the choice set. For this purpose, we use the patient model, which we know to take into account patients' preferences. After calculating the probabilities of choice for all the patients, we are able to average the probabilities by hospital, and to predict the total number of patients that will choose it, i.e., its predicted volume of activity. Knowing that the volume is related to the clinical quality, we can predict also the total mortality by hospital and, summing up all of them, the total mortality associated to the choice set. Since different choice sets will give birth to different patients' behaviours, which in turn will affect volumes and will end up in distinct outcomes, we are able to make comparisons among outcomes resulting from different choice sets.

By using this solution procedure, we keep the decision maker perspective: the actions of the patients are seen from a strategical perspective, and the major interest remain the clinical quality. However, we are enriching the decision maker point of view with the consideration of patients' preferences, which allow a more precise prediction of how patients will answer to distinct choice sets. As already done in Section 3.4, in order to focus specifically on the quality of the healthcare services and on a strategical level, we are not inserting any consideration about costs. The solution procedure presented is referred to the annual allocation of volume, and can be replicated for each year of interest.

The solution approach (Figure 5.3) consists of three phases.

Phase 1 During Phase 1, we explore all the strategical choice sets that are available, in terms of combinations of opening/closures of regional hospitals. We do not explore the different volumes of activity that could be attributed to hospitals upon the policy maker model. In fact, we know that patients will act only considering which hospitals are present in the configuration, together with their distance and past performance. The choice set is, from the patient perspective, the unique long-lasting decision made by the decision maker.

It should be noticed that, if the number of provinces and hospitals in the territory increases, the number of possible strategical choice sets increases as well. Additional criteria to decrease it could be adopted, in order to (i) reduce the complexity in terms of computation time needed to reach the optimal solution; (ii) make the solution more realistic, by avoiding to suggest closure of hospitals that are not feasible in

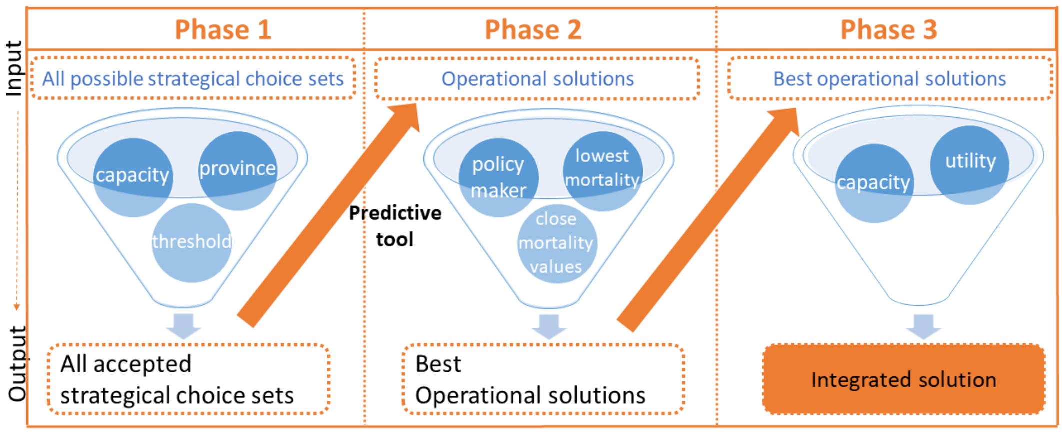


Fig. 5.3 Description of the algorithm of the fully-integrated approach

the short terms due to organizational or social reasons. In particular, we opted for a criterion that impose hospitals to remain open if they have performed more than 50 interventions during the previous year.

All the possible strategical choice sets go through a selection process, where three decision maker's constraints are tested. First of all, it is verified if the hospitals in the choice set performed the previous year at least the T threshold volume advocated by international guidelines. Secondly, the hospital is considered to have a limited capacity. Third, the sum of the capacities of the open hospitals within a province needs to be sufficient to satisfy the provincial request of interventions. This last constraint takes the extreme case of a null hospitality threshold, where each province provides service to its own inhabitants and patients' mobility among provinces is excluded. Once all the *accepted* strategical choice sets are determined, Phase 2 begins.

Phase 2 For each strategical choice set, the patient model is used to calculate the expected probability that each patient will choose each alternative of the choice set. By using the predictive tool of the conditional logit (equation 5.10), the average probabilities of the population are assigned to the hospitals as percentages of population that will choose them. In this way, we are able to predict the number of patients that will choose each hospital, which generates the *operational solution*, i.e. the allocation of volumes of activity among hospitals. We then calculate the mortality associated to each *operational solution*. From all the solutions, we select the *operational solutions* that respect three conditions:

1. the solutions respect two decision maker's constraints: (i) hospital volumes resulting from patient choices have to be greater than the T threshold of interventions; (ii) operational solutions have to respect hospitals capacity. As the choices of patients might have altered these two quality requirements, we allow for a partial relaxation of the involved constraints through the use of a discrepancy percentage $\alpha_d\%$. In fact, we reckon the phenomenon could need to be addressed in a longer period of time;

2. the solutions are associated with the three lowest integer mortality values;

3. the solutions associated with the second and third lowest mortality values do not exceed the first lowest mortality value of more than $\alpha_m\%$. If this condition is not met, we keep only the solutions associated to the first lowest mortality value.

The output of Phase 2 will be a set of *operational solutions*, which will be from a minimum of one (if the two solutions with the second and third lowest mortality values exceed the first mortality value of more than $\alpha_m\%$, and there is only one solution with the minimum mortality value), up to more than three, if multiple solutions share the same lowest mortality value. If there are multiple solutions, Phase 3 is needed in order to decide which one should be chosen.

Phase 3 Phase 3 consists in the skim process of the selected *operational solutions*. The first criteria of the skim process is based on the comparison among the volumes calculated from the *operational solution* and the *strategical solution*. In fact, each strategical choice set can be used as an input of both the decision maker model and the patient model. Even if the strategical choice set is the same, the output of the two models differ, since the *strategical solution* originates from the minimization of the mortality, while the *operational solution* originates from patient preferences. We are interested in analysing this difference in terms of capacity required to hospitals. More specifically, we are concerned about the impact of patients' distribution among hospitals. Hospitals capacity is decided in advance by the decision maker, with the aim of helping hospitals to adapt their structure and organization to the required volumes. The policy maker makes the decision as a consequence of the *strategical solution*. Hence, we prefer *operational solutions* whose volumes are less discrepant from the ones allocated in the *strategical solution*. In this way, facilities will be less unprepared to face the actual demand, since it will be similar to the one planned by the policy maker. Equation (5.19) describes how the *capacity discrepancy index* (*CDI*), is calculated:

$$CDI = \frac{1}{\sum_{n=1}^{N} J_n} \cdot \sum_{n=1}^{N} \sum_{j=1}^{J_n} \frac{abs(x_{j_n} - \sum_{n=1}^{N} \sum_{i=1}^{I_n} p_{i_n,j_n})}{x_{j_n}} \tag{5.19}$$

We average across all the hospitals of all the provinces the difference (absolute value) among the volumes that are planned by the decision maker model (x_{j_n}) and the volumes that result from patient preferences ($\sum_{n=1}^{N} \sum_{i=1}^{I_n} p_{i_n,j_n}$), and we normalize this difference by the volumes planned by the decision maker model. We then select the *operational solution* with the lowest capacity discrepancy index.

If multiple *operational solutions* assume the same minimum capacity discrepancy index value, a further skim-process criteria is represented by patients' utility. We assume that the decision maker is interested in maximizing the satisfaction perceived by patients, by minimizing the difference existing among expected and actual utility. The expected utility is defined through the information that patients have available i.e., past information. The actual utility is calculated by knowing what happens in the reality. In particular, we want to see if the predicted volumes would have been the same by calculating utility not with the past performance (the source of information used by patients), but rather with the actual performance (that depends on the choices of all the patients). The change in volumes can be globally interpreted as a change in patients' choices, i.e., a measure of utility discrepancy. Equation (5.20) describes how the *utility discrepancy index* (*UDI*), is calculated:

$$UDI = \frac{1}{\sum_{n=1}^{N} J_n} \cdot$$

$$\cdot \sum_{n=1}^{N} \sum_{j=1}^{J_n} \frac{abs(\sum_{n=1}^{N} \sum_{i=1}^{I_n} p_{i_n,j_n}(x_{j_n,t-1}) - \sum_{n=1}^{N} \sum_{i=1}^{I_n} p_{i_n,j_n}(x_{j_n,t}))}{\sum_{n=1}^{N} \sum_{i=1}^{I_n} p_{i_n,j_n}(x_{j_n,t-1})} \tag{5.20}$$

(5.21)

We average, across all the hospitals of all the provinces, the difference (absolute value) among two volumes. The first volumes are the results from patient preferences ($\sum_{n=1}^{N} \sum_{i=1}^{I_n} p_{i_n,j_n}$) using the past information about hospital performance ($x_{j_n,t-1}$). The second volumes are the results from patient preferences using the present information about hospital performance ($x_{j_n,t}$). The difference is then normalized for the volumes resulting from patient preferences using the past information about hospital performance. Ultimately, the *strategical-operational* solution with the lowest UDI will be chosen for the territorial configuration.

5.1.5 Comparative analysis

The second objective of this chapter is to investigate the difference among solutions that are identified as optimal by considering separately the two perspectives, or by partially/fully integrating them. In order to do so, a framework of key indicators was created. All of them are devoted to deepen pros and cons of the proposed configurations. The indicators are based on three main categories, referring to three aspects upon which the configurations can be evaluated: mortality, volume, distance. All the indicators are designed to be calculated for each year, hence, in the following, the index t will not be used.

The first category is mortality, which remains the most relevant outcome. Besides the total mortality, we also calculated the values corresponding to the other three objective functions presented in Chapter 3:

- Total mortality M_1:

$$M_1 = \sum_{n=1}^{N} \sum_{j=1}^{J_n} x_{j_n} \cdot m(x_{j_n}); \tag{5.22}$$

- Average mortality M_2:

$$M_2 = \frac{1}{\sum_{n=1}^{N} J_n} \sum_{n=1}^{N} \sum_{j=1}^{J_n} m(x_{j_n}); \tag{5.23}$$

- Mortality variance M_3:

$$M_3 = \frac{1}{\sum_{n=1}^{N} J_n} \sum_{n=1}^{N} \sum_{j=1}^{J_n} [\mu - m(x_{j_n})]^2; \tag{5.24}$$

- Mortality range M_4:

$$M_4 = [\max_{n,j_n}(m(x_{j_n})) - \min_{n,j_n}(m(x_{j_n}))]. \tag{5.25}$$

For all the four indexes related to mortality, low values will act as a signal of quality.

The second category is volume, whose consideration has two advantages. First, it confirms what the mortality highlights, since the two indicators are linked through the volume-outcome association. Second, it sheds some light on what patients look at. In fact, compared to mortality, information on volumes are easier to be obtained by patients. This type of data does not need to be adjusted for patient characteristics, and it does not require additional explanation that are harder to vehicle. Moreover, characteristics that are publicly visible, like the size of the hospital structures, are a good proxy for volumes. Hence, we can realistically assume that patients are aware of the average number of people that visit hospitals (i.e., their volume of activity). Through this data, they can infer information about surgeons' experience. Within this category, we analyse four indexes:

- Volume threshold V_1: the percentage of hospitals that perform more than the threshold T of interventions. Since the threshold T is suggested by international guidelines, we would expect V_1 to be automatically set to 100%. However, this is not the case in the reality, hence it needs to be monitored. The indicator is calculated as:

$$V_1 = 100 * \frac{\sum_{n=1}^{N} \sum_{j=1}^{J_n} \lfloor \frac{x_{j_n}}{T} \rfloor}{\sum_{n=1}^{N} J_n}; \tag{5.26}$$

- Capacity ceiling V_2: the percentage of hospitals that perform a number of intervention that is lower than their capacity. V_2 acts as a proxy of the organizational quality. We prefer those configurations that guarantee all the hospitals to host no more patients than they are able to. It is calculated as:

$$V_2 = 100 * \left(1 - \frac{\sum_{n=1}^{N} \sum_{j=1}^{J_n} \lfloor \frac{x_{j_n}}{cap_{j_n}} \rfloor}{\sum_{n=1}^{N} J_n} \right). \tag{5.27}$$

These two indexes will clearly have maximum value of 100% in the policy maker model (and, hence, in the quasi-integrated approach), where both are included in the constraints. Once the decision maker solution is set as the first best, it is interesting to explore the discrepancy among the other solutions that leave free space to patient choice. In fact, since patients do not have a complete knowledge of the volumes that origin from their individual choices, they could undermine the quality (i.e., V_1) and the organization (i.e., V_2) of several hospitals.

- Mean volume between hospitals V_3: the mean number of interventions among hospitals. It is calculated as:

$$V_3 = \frac{1}{\sum_{n=1}^{N} J_n} \sum_{n=1}^{N} \sum_{j=1}^{J_n} x_{j_n}. \tag{5.28}$$

  This index will help in understanding how volumes are spread among hospitals. It could be more appropriate to look at quantiles, but preliminary analyses show that the conclusions are the same (i.e., in the configuration with the highest mean value, all the quantiles values will be the highest as well). High values of V_3 suggest a good level of surgeons' experience.

- Mean volume between patients V_4: the weighted average of interventions. It is calculated as:

$$V_4 = \frac{1}{\sum_{n=1}^{N} I_n} \sum_{n=1}^{N} \sum_{j=1}^{J_n} x_{j_n} \cdot \sum_{n=1}^{N} \sum_{j=1}^{J_n} x_{j_n}. \tag{5.29}$$

  Differently from V_3, this index is a weighted mean that accounts for some data points (those corresponding to higher volumes) to contribute more than others. The idea is that V_4 represents the patients' perspectives on the volumes performed at a regional level. If we assume that patients talk to each other, a hospital that performs 100 interventions will translate into 100 patients telling that their hospital has performed 100 interventions. On the contrary, the presence of a small hospital that perform only 10 interventions will transform into only 10 patients telling that their hospital has performed a low volume. As a consequence, no matter how many small hospitals are open in the territory, patients will perceive more the presence of big hospitals, and this will affect the positive/negative impression that they have about the regional performance. Hence, this index is quite valuable for our analysis since it represents the information that patients have, thanks to word of mouth, about the regional performance.

The third aspect is distance, which allows us to enlighten pros and cons as seen by patients' perspective. Within this category, we studied two indexes:

- Mean travelled distance D_1:

$$D_1 = \frac{1}{\sum_{n=1}^{N} I_n} \sum_{n=1}^{N} \sum_{j=1}^{J_n} \sum_{n=1}^{N} \sum_{i=1}^{I_n} d_{i_n,j_n} \cdot p_{i_n,j_n}. \tag{5.30}$$

  This index enlightens the average amount of kilometres travelled by patients.

- Mean distance from the closest hospital D_2:

$$D_2 = \frac{1}{\sum_{n=1}^{N} I_n} \sum_{n=1}^{N} \sum_{j=1}^{J_n} \sum_{n=1}^{N} \sum_{i=1}^{I_n} d_{i_n,j_n} \cdot min_{i_n,j_n}. \tag{5.31}$$

  This index reveals the lowest distance that patients need to travel in order to be treated. It deserves attention for two reasons: (i) it gives a measure of the welfare guaranteed by the decision maker to the population; (ii) if compared to D_1, it helps to understand the choice made by patients, in terms of the difference among what they could have done (D_2) and what they have decided to do (D_1).

For both the indicators of distance, low values will be preferred.

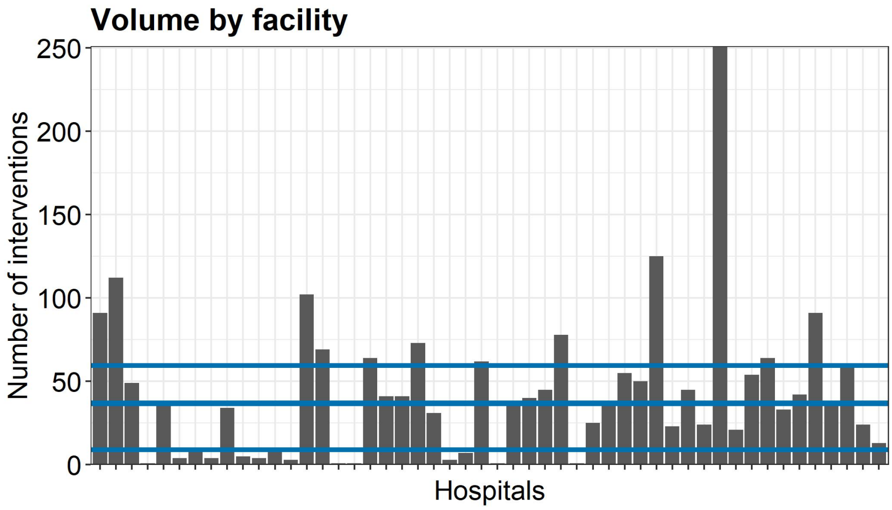


Fig. 5.4 Number of colon cancer interventions performed in hospitals in Piedmont, 2014

5.2 Case study

The solutions that are identified as optimal by considering separately the two perspectives, or by partially/fully integrating them, were tested and compared on the real case study used in Chapter 4, which concerns the regional distribution in Piedmont of hospitals for colon cancer surgery. In particular, we focused on patients that received a colon cancer surgery in 2014, which accounted for 1405 people. The regional context perfectly presents the prerequisites for our models and approaches to be applied, both for the relevance of the illness and for the scattering of small hospitals in the territory (Figure 5.4 reports the volumes by facility; lines identify quartiles).

The following sections will describe the configurations resulting from the different perspectives. The results from comparative analysis will be presented in Section 5.2.5.

5.2.1 Policy maker solution

The starting point of the application of the policy maker model is the actual territorial configuration: the available alternatives in the choice set are all the hospitals that operated colon cancer surgery in 2013. By considering the possibility for the decision maker of closing some of these hospitals, we aim to evaluate which changes in the volume distribution (and, as a consequence, in the mortality) in 2014 there would have been if the decision maker model had been applied.

In order to reduce the computational complexity, we consider the special case of each province providing service to its own inhabitants, i.e., the hospitality threshold is set to zero ($\delta = 0$). Hence, we exclude patients' mobility among provinces. By doing so, we can separately treat each province, i.e., decompose the complete model in N sub models, one per province. Moreover, we do not insert any consideration about costs (constraints (5.6)), which mostly deserve attention on a operational level.

As for the volume:

1. related to the upper bound (i.e., capacity), we assumed that hospitals can perform up to two times the volume performed during the previous year. This assumption is reasonable as we are deciding on a strategical level, where hospitals can be thought to have the possibility to deeply change their internal organization, e.g., hiring more personnel, or using operating rooms for other interventions, etc.

2. related to the lower bound, we allowed hospitals to remain open if they have performed in the previous year $T/2$ interventions. We left the original lower bound at T only for the province of Turin. In fact, in the territory there is a high concentration of small hospitals, and in some provinces (e.g., AL) the hospitals that satisfy the constraint on the T threshold would not have enough capacity to host all their provincial demand.

3. if hospitals have performed more than T interventions in the previous year (whatever province they belong to), we force them to remain open. In fact, we reckon that it is more realistic to close hospitals that are recognized to perform not enough interventions, rather than those that have respected the international guidelines on volumes.

Figure 5.5 shows the *strategical solution* resulting from the solution of the N decision maker models, and compares it to the actual solution. The suggested configuration is pretty obvious: the minimization of the total mortality drives to the full concentration of interventions in few hospitals. Since the hospitality threshold is set to zero, it forces the presence of at least one facility per province, as it happens in the provinces of Asti, Biella, Verbania and Vercelli. Additional facilities are present

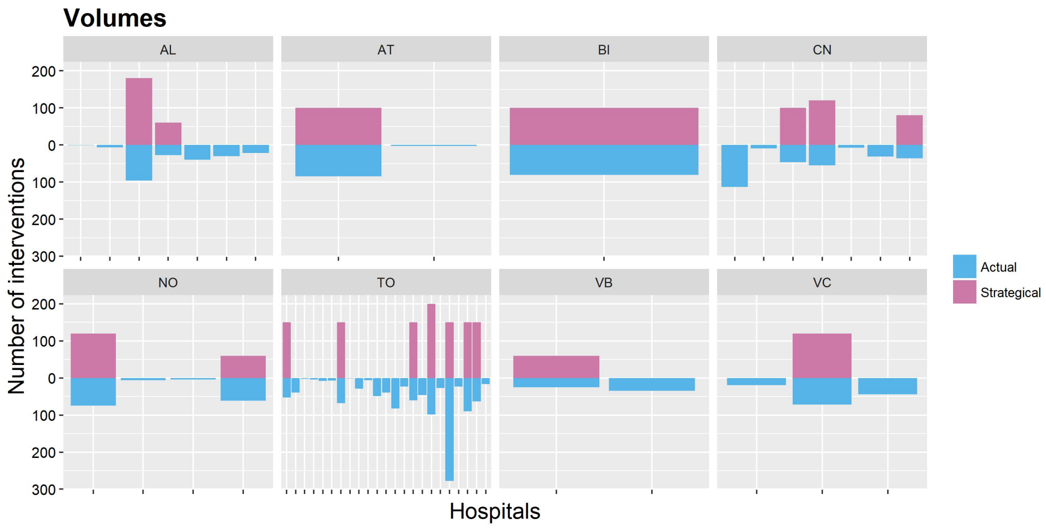


Fig. 5.5 Comparison among volume distributions of actual and strategical solutions

in the provinces of Alessandria, Cuneo, Novara and Torino, where the capacity of the biggest hospital is not sufficient to cover the whole population demand. Because of the previous performance information, the model assigns the provincial volume to the hospitals that performed best during the previous year.

However, the *strategical solution* should not be confused with what is going to happen in the reality. In fact, since patients are free to choose any hospital in the regional choice set, their choices could cause changes in the volume distribution (with respect to the one predicted from the *strategical solution*) and, as a consequence, in the clinical outcome as it will be shown in Section 5.2.2.

5.2.2 Patient solution

The patient model can be used to predict the patients' behaviour since it calculates, for each patient, the probabilities of choosing each hospital. Similarly to what we presented in Section 4.3.3, we applied the conditional logit to all the patients from Piedmont that received colon cancer surgery from 2004 to 2013, to obtain the values for the coefficients β_{d1}, β_{kd2}, β_{q1}, β_{kq2}. These coefficients were then used in a predictive model, focusing on the group of patients that needed colon surgery in 2014.

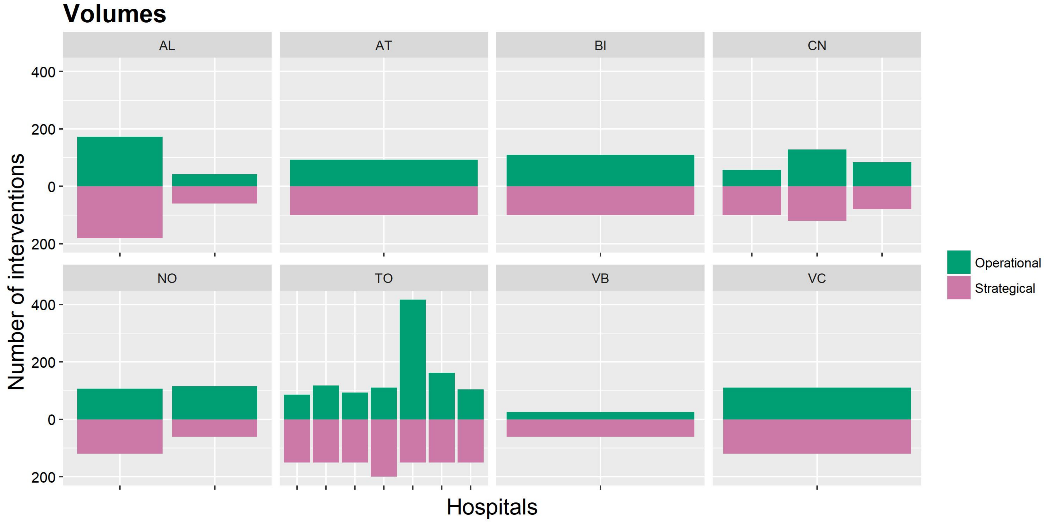


Fig. 5.6 Comparison among outcomes of actual and strategical solutions

We applied the predictive model to the choice set resulting from the *strategical solution*. The volumes resulting from patients' decisions represent the *operational solution*, and are shown in Figure 5.6, together with the ones originated from the *strategical solution*. It should be noticed that, since the decision maker model does not focus on the single patients actions, we cannot compare *operational solutions* and *strategical solutions* at the individual level. As an example, even if in the province of Alessandria the resulting volumes from the policy maker and the patient models are very similar, we cannot state that patients *behave as expected*, because the result of the policy maker model had no expected volumes at the single individual level.

The only plan on individual level that the decision maker proposes is, for each province, that all the patients are treated in their provincial hospitals. By looking at the *operational solution* in terms of probabilities of choice, we see that the average probability of patients to choose any hospital in the same province is 17%, while it decreases to 0.9% for any hospital outside of the provincial borders. Moreover, if we restrict the sample only to patients that have a unique hospital in their province (i.e., patients from the provinces of Asti, Biella, Verbania and Vercelli), the average probability of choosing the unique open hospital in the province becomes 64%. As a consequence, we can infer that these patients, mainly due to the disutility for distance, will be much more prone to choose their closest hospital. On the other hand, the closure of multiple hospitals, by increasing the volumes of the open hospital, increases the utility that patients perceive from this choice.

From Figure 5.6 we can see that, in most of the hospitals, the aggregation of patients' choices does not mirror planned volumes. This discrepancy raises concerns related to organizational issues. While the increase/decrease of volume (compared to the number of interventions performed in 2013) required by the decision maker represents a strategical decision that is communicated in advance, the variation of volumes caused by patients' choices does not allow facilities to be prepared to it. As an example, a hospital in Torino (San Giovanni Battista Molinette) would receive by the decision maker model the guideline of performing 150 interventions, which represents 27% of its capacity (it performed 278 interventions in 2013, thus we assume its capacity to be 556). We can realistically imagine the operational reallocation of the remaining 73% of its capacity, to other duties, interventions, etc. However, the following year, due to patients' choices, it would unexpectedly experience a total demand of 389 interventions. We are not able to quantify the consequences of this overload and of the organizational changes faced by hospitals, but we can assume that also clinical outcomes would be negatively affected by crowding and personnel stress.

By comparing at the aggregated level the *operational solution* and the *strategical solution*, we observe that 35% of patients make a choice that differs from what had been planned by the policy maker. By using the patient model, we performed also an additional analysis meant to understand how this patients' behaviour could be influenced by the availability of information on the choice set. Specifically, we calculated how many patients, had they known the decision maker plan in terms of volumes (not only of opening/closures), would have made a choice different from the one thought for them by the policy maker. To this purpose, we again predicted patients' behaviour using the patient model, but this time applying the hospital characteristics that result from the *strategical solution* (in terms of opening and volumes assigned). The interesting result is that, in this case, only 17% of patients would have made a different choice. This result promotes further reflections on the integration between the decision maker's and patients' perspectives. In fact, the difference among 35% and 17% can be attributed to the patients' knowledge of the overall territorial distribution of volumes. Hence, the policy maker needs to address the problem of making patients fully aware of his/her regional plan.

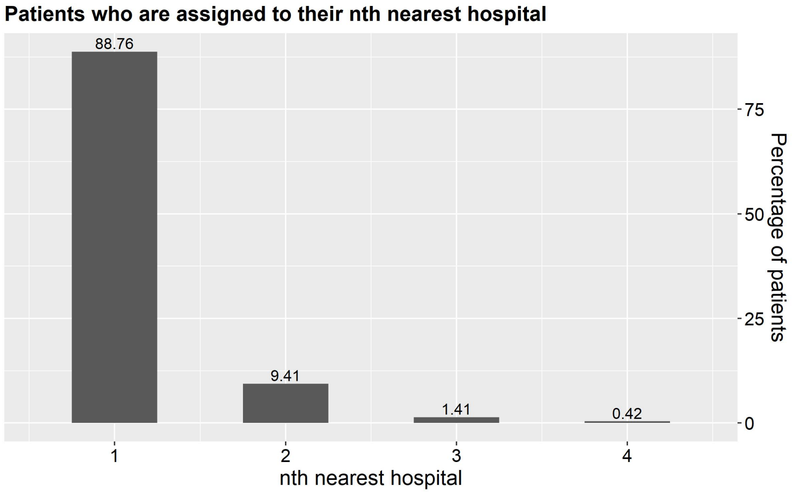


Fig. 5.7 Patients who are assigned to their nth nearest hospital, quasi-integrated approach

5.2.3 Quasi-integrated solution

The application of the quasi-integrated approach to the case study simply consists in applying the policy maker model as in Section 5.2.1, but focusing on each patient action, and guaranteeing to all the patients that their constraint about the willingness to travel is respected. For the colon cancer case study we are dealing with, the *strategical solution* proposed in Figure 5.5 respects the new constraints of the quasi-integrated model when using α_{wtt} equal to 11%, hence the policy maker model and the quasi-integrated approach will have solutions with the same characteristics at a hospital level.

However, it is interesting to look at the solution of the quasi-integrated approach from an individual perspective. In fact, the outcome of the quasi-integrated approach includes planning decisions not only on hospital volumes, but mostly on individual choices. We then investigate the distance that patients are required to travel, based on the solution established by the decision maker. We can analyse this aspect from two points of view, shown in Figure 5.7 and Figure 5.8. Figure 5.7 shows the percentage of patients that are assigned to their nth nearest hospital. It can be seen how the decision maker plans to direct almost all the patients (88.76%) to their closest hospital. Less than 12% of patients is thought to choose a further hospital, and less than 1% of patients is supposed to travel further than the third closest hospital.

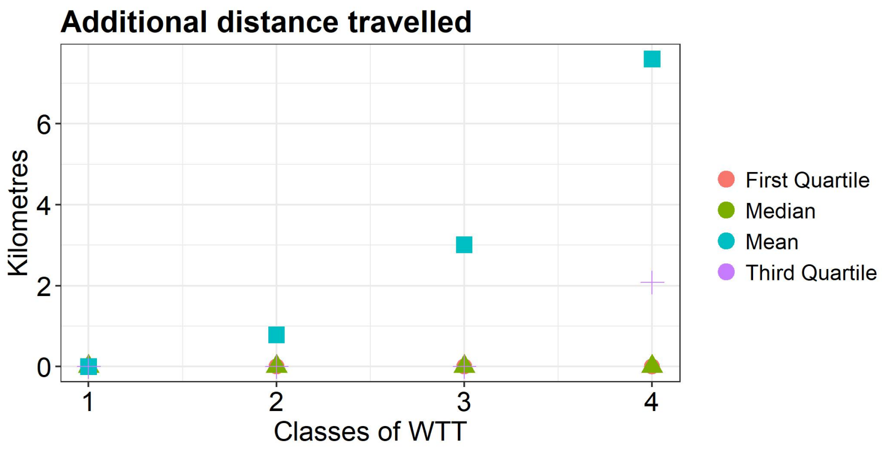


Fig. 5.8 Additional distance travelled, quasi-integrated approach

What is most interesting is to distinguish the additional distance travelled by patients based on their willingness to travel, as shown in Figure 5.8. Classes have been defined based on quartiles of the population values for the willingness to travel. Two things can be noticed: for all the classes of willingness to travel but the highest, more than 75% of patients does not travel any additional kilometre (all the coloured dots are overlapping each other since they are all on the same y-axis position); there is a small number of patients that travels an additional distance, and this distance drastically decreases when the willingness to travel decreases. In other words, in very few cases, the policy maker proposes to patients with a low willingness to travel to go further than the closest hospital, and, in those few cases, they are guided to other hospitals that are, on average, only few kilometres further.

The quasi-integrated approach allows to reach a solution that is optimal for the population health (i.e., it reduces mortality), and that takes into account what patients are not willing to do by bounding from above the maximum travelled distance. Nonetheless, the implicit assumption is that patients will behave exactly as the policy maker has decided for them. How much this assumption is realistic strictly depends on the communication tools used by the policy maker to convey the messages to the population. If patients are all well informed about (i) the choice that has been thought to be the best for them and (ii) the health conditions that this choice guarantees to the whole population, they will more likely adhere to the decision maker's guidelines.

5.2.4 Fully-integrated solution

The fully-integrated approach overcomes the difficulties related to the communication among patients and decision makers, since the decision maker takes charge of predicting patients' behaviours and adapting the planning decisions to their expected answer. In order to apply this approach, we set both α_m (i.e., the distance among the three mortality values identified in Phase 2) and α_d (i.e., the discrepancy among the decision maker's constraints tested in Phase 2) to 10. For all the considerations on hospitals and volumes, see Section 5.2.1.

As described in Section 5.1.4, in Phase 1, we explore all the possible combinations of openings and closures of hospitals. In order to respect the selection criteria on the volume threshold, we excluded the choice sets with hospitals that performed the previous year less then T interventions. However, since in several provinces (e.g., Biella) the number of remaining hospitals would not be sufficient to cover the provincial demand, we relaxed the constraint, for all the provinces but Turin (where a high number of hospitals respected the T threshold), to include all the hospitals that performed more than $T/2$ interventions during the previous year. At this point, we excluded the choice sets with a capacity (at a provincial level) smaller than the number of required interventions. The output of Phase 1 is 41,580 strategical choice sets.

At the beginning of Phase 2, we generate the *operational solutions*, i.e., the predictions of the percentages of the population choosing each available hospital. For each *operational solution*, we calculate the mortality associated to it. Figure 5.9 shows the distribution of mortality values for all the *operational solutions*. Depending on the strategical choice set that originates the *operational solution*, hospital volumes resulting from patients' choices can widely differ, and, as a consequence, mortality values vary as well.

By checking the constraints on the capacity and on the threshold T, and allowing for a discrepancy of 10%, we select only the 3% of operational solutions.

Since the second and third lowest mortality values do not exceed the first mortality value of more than 10%, and there are no multiple solutions with the lowest three integer mortality values, the output of Phase 2 is composed by three operational solutions (and, hence, three strategical choice sets), with a mortality ranging from 76 to 77. This gap represents a variation of less than 2%, hence, from a clinical

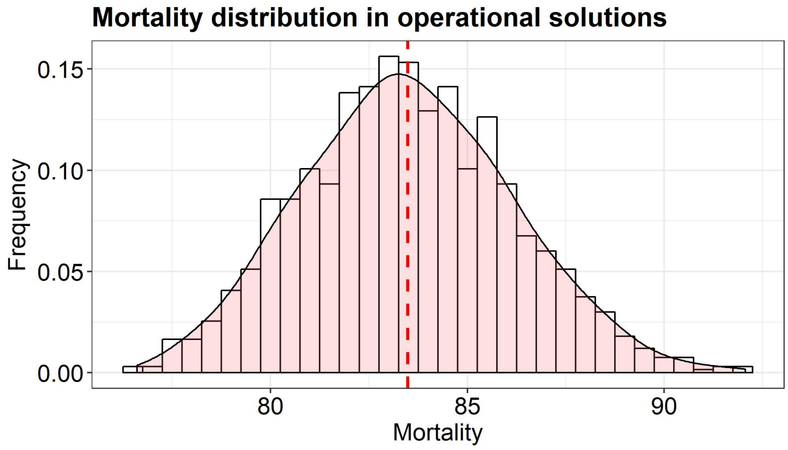


Fig. 5.9 Mortality for operational solutions

perspective, the group of solutions are interchangeable. Phase 3, by adding additional criteria to define quality, has the objective of identifying the solution to be preferred.

In order to start the skim process of Phase 3, we run the decision maker model for the strategical choice sets that are the outputs of Phase 2. For each strategical choice set, we match the *strategical* and the *operational solution* that corresponds to it, which we define as *matched*.

The first analysis of Phase 3 is based on the capacity required by the *matched* solutions, by computing their capacity discrepancy index, with equation (5.19).

Since there is only one *operational solution* that reaches the lowest value for capacity discrepancy (which is also the same solution that has the lowest mortality value), we identify it as the optimal solution of the integrated perspective, namely the *integrated solution*.

5.2.5 Comparative analysis

Table 5.4 reports the value of the key performance indicators for each solution approach and for the actual situation.

Since the most important outcome of our analysis remains the total mortality (M_1), Figure 5.10 reports the values of mortality resulting from the (i) actual configuration, (ii) policy maker model (that, as for mortality, has the same performance of the quasi-integrated approach), (iii) patient model, (iv) fully-integrated approach. They

Table 5.4 Key performance indicators for the comparative analysis

		Solution				
		Actual	Policy maker	Patient	Quasi- Integ	Fully- Integ
Mortality	M_1	96	74	80	74	76
	M_2	7.97%	1.34%	1.38%	1.34%	2.77%
	M_3	46.72	4.26	4.04	4.26	3.26
	M_4	24.1	7.91	6.93	7.91	6.68
Volume	V_1	31%	100%	84%	100%	94%
	V_2	100%	100%	100%	100%	95%
	V_3	44	116	112	116	133
	V_4	86	129	161	129	198
Distance	D_1	14.98	*NA*	18.27	16.37	19.9
	D_2	7.2	14.57	14.57	14.57	15.7

are all comparable, since they are a prediction calculated by using the mortality curves of the volume-outcome association.

The best solution in terms of mortality is the *strategical solution*, which we expected since it is the only solution that comes from the objective function of minimizing the mortality itself.

The first difference emerges between the mortality resulting from the *actual solution* and the *strategical solution*. The combination of a lower number of hospitals (hence performing higher volumes) and the closure of facilities with low performance coefficients, causes the strategical solution to lead to an improvement of the 23% of mortality, which decreases from 96 to 74.

The comparison between the *strategical solution* and the *operational solution* confirms that the strategic plan developed by the policy maker conducts to a better outcome, whereas the free choice of patients can lead to a lower clinical quality. However, it also shows that both the *strategical solution* and the *operational solution* present a huge improvement if compared to the *actual solution*. Two causes can be identified for this result: (i) the decrease in the number of available hospitals (due to the policy maker solution), even if patients' free choice is allowed, globally increases the volume of interventions and, as a consequence, the quality; (ii) hospitals that are available to be chosen have been selected through the performance coefficient, hence they will better perform any volume of activity. All in all, through the decision on the choice set made by the policy maker, patients are forced to choose more wisely.

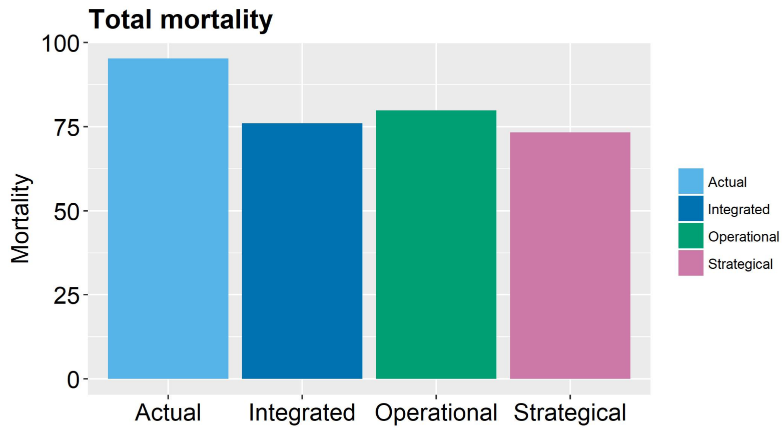


Fig. 5.10 Comparison among outcomes of actual, strategical, operational and fully-integrated solutions

Interestingly, the mortality resulting from the *fully integrated solution* is higher than the one from the *strategical solution*, but lower than the one from the *operational solution*. In fact, the *fully integrated solution* represents a choice that we would have not defined as optimal if looking from the policy maker perspective, but that comes to be optimal when we check for patients' answer. Of course, the first best from a theoretical perspective remains the solution resulting from the policy maker model. But if we aim to guarantee to the policy maker that the solution will be realistically adopted, the *fully integrated solution* becomes the optimal one.

As for the other indicators of mortality, they all highlight different aspects of the quality of the solutions. The *actual solution* remains the only one that has always worse level of quality. As for the average mortality (M_2), the *strategical solution* and the *quasi integrated solution* have the lowest value. The *fully integrated solution* has the best value of both mortality variance (M_3) and mortality range (M_4).

As far as volume is concerned, the volume threshold (V_1) signals the absence of a policy of specialization in the reality. In fact, in the *actual solution* only 31% of hospitals perform more than T interventions. Since the policy maker model has one constraint dedicated to force the threshold T as lower bound of the volume of activity, we already knew V_1 to assume maximum value in the *strategical solution* and in the *quasi integrated solution*. This indicator decreases for the *fully integrated solution* (where we allowed for a 10% discrepancy from the policy maker constraint), and even more in the *operational solution*.

Similar conclusions on the policy maker model and the quasi-integrated approach can be drawn for the indicator on the capacity ceiling (V_2), since a constraint forces all the hospitals to respect their capacity. However, V_2 assumes high value (100%) also in the *actual solution*. This can be explained by the absence of a reconfiguration policy, which makes the demand to be more or less stable from one year to the following. The *operational solution* has maximum value as well, which can be explained by the fact that the choice set that is available to patients has been established by the policy maker. As for the fully integrated solution approach, the skim process based on the minimization of the capacity discrepancy index positively affects the result of the indicator V_2.

Related to the mean volume between hospitals (V_3) and between patients (V_4), it is interesting to notice that, compared to the policy maker model, the patient model drives to a better value for V_4, but a worse value for V_3. This finding confirms our interpretation of the indicator V_4, which represents what patients are aware of thanks to the word of mouth. Interestingly, the *fully integrated solution* appears as optimal for both the indicators. This result is reassuring as for the probability of patients to be compliant.

Eventually, the two indicators on distance reveal some information on how patients behave vs. how patients could behave. For instance, even if patients have the closest hospital, on average, only 7 kilometres far, they travel on average more than two times this value to reach the chosen hospital. Interestingly, the distance to the closest hospital guaranteed by the *strategical solution*, even if much higher than the one of the *actual solution*, is lower compared to what patients on average do. Hence, there exist a rationale for the policy maker to demand to patients to travel, in general, more. However, since D_1 and D_2 are formulated as averages, attention should be paid to who are the patients that, in reality, are travelling more.

All in all, considering the output on mortality, two extreme cases can be identified: the case resulting from the *actual solution*, which acts as the worst case, and the case from the *strategical solution* (together with the *quasi-integrated solution*), which represents the best case. Given that the first one will be changed anyway, and the second one has been said to be effective only on a theoretical level, attention should be paid to the other two solutions. Both *operational solution* and *fully integrated solution* globally enhance population health, but the *fully integrated solution* reaches better quality in terms of total mortality. By looking to all the other indicators, the *fully integrated solution* has 6 out of 9 indicators with better values. Beyond the indicators we built, there are multiple managerial factors that make the *fully integrated solution* more desiderable. In fact, by predicting patient behaviour, hospitals can be prepared to face patients' answer, even though it is different from the policy maker plan. This represents a huge added value, since it avoids overloads of capacity, burn out of medical staff, inefficiency in the surgery processes.

5.3 Policy Implications

Behind the findings in terms of the key performance indicators, the consequences that the integrated approach has must be carefully evaluated. Since it entails a change on the whole planning process, these consequences regard all the stakeholders of the healthcare system.

- **The policy maker**. The use of the integrated approach translates into a change in his/her decisional process, which brings several advantages.

  First of all, the new decisional process is built on principles of evidence based medicine (EBM), rather than on political or only managerial concerns. This means that the rationale for the required structural and behavioural changes is supported by the scientific evidence. Therefore, the policy maker is less subjected to critics and scepticism for the reasons that have guided the planning choices.

  Moreover, since the integrated approach includes the prediction of patient behaviour, it increases the probability of patient adherence, and, as a consequence, the likelihood of a successful implementation of the planning decisions.

  Most importantly, given the main objective of the policy maker, i.e., the maximization of the population health conditions, the new decisional process ensures to target it.

  However, several criticalities need to be tackled by the policy maker. First of all, a decisional process that is based on the scientific evidence, calls for a high quality of the evidence itself, both in terms of data used and of their analysis. For this reason, more attention must be paid to three aspects. The first one is the process of data collection within hospitals. Physicians have to receive specific training, in order to avoid measurement bias due to negligence, lack of understanding, or low priority given to this task of their job. The second aspect is the commissioning of the analysis of clinical data to expert researchers. The third aspect, which resumes both of the previous two, is the process of audit. The process of audit consists in verifying the quality of the data, after the data have been analysed. As an example, in the PNE system, after the presentation of data on-line, providers and physicians have the possibility to raise questions and concerns on the performance scores they have received. The debate helps the researchers to tune the methodologies used for the analysis of data (e.g., to discover the most relevant factors to be considered for the risk-adjustment techniques), and, at the same time, it helps physicians and providers to understand the existing scope for improvement. Handling these three aspects (physicians training, commissioning to data analysts, audit process) will favour the integrated approach to be updated, according to both changes in the practice and in the evidence as well.

Another controversial aspect that is faced by the policy maker is to guarantee the correct implementation of the changes suggested by the evidence, e.g., the reallocation of volumes among hospitals. In this sense, the creation of incentives acts as a possible support to promote the change, even if attention must be payed to the risk of the so called *gaming*, i.e., opportunistic behaviours on behalf of the actors involved. Examples of incentives are the volume threshold and the pay for performance. Imposing a threshold of interventions that need to be performed is useful to make the providers aware of the importance of the volume-outcome association. Nonetheless, this incentive could lead to weird increases in volumes above the threshold, which could hide over-treated cases that receive the intervention exactly to overcome the threshold.

In the following points, we will describe some of the consequences of the integrated approach that are faced by the shareholders of the healthcare sector, other than the policy maker. However, all these other factors are somehow related to the policy maker's perspective, as the policy maker is the actor with the objective of taking care of the whole population.

- **Providers**. The policy built on the use of the volume-outcome association has, as a consequence, the reorganization of the structures that provide healthcare services. Hospitals need to adapt their spaces and staff to the new volumes of activity advocated by the quality standard. Hence, providers are required to think of new management strategies that guarantee efficiency within the new organization. Nevertheless, the increase of volumes can be exploited to reach economies of scale. Moreover, since the concentration of procedures is related to a decrease in mortality, and quality becomes the aspect on which providers are evaluated, the reorganization assures higher quality scores.

- **Physicians**. The change in the allocation of volumes among hospitals will have effect also on the organization of the work of the medical staff. Surely this change will require attention by the human resources department, in order to avoid burn-out, dissatisfaction or not desired transfers caused by hospitals closure.

  However, the most important change that is stimulated by the managerial use of the volume-outcome association is the evaluation of the medical staff based on clinical outcomes. By using discharge records that include the identification of the surgeon, it will be possible to map the surgeon performance over time. Even if there exists the risk for physicians to feel stressed by perceiving the evaluation as a judgement, we reckon there is a high probability of feeling rewarded for the improvement in their job.

- **Researchers**. The priority given to the creation of evidence based medicine increases the importance of the work made by researchers in the epidemiological field. More emphasis has to be placed then on multiple aspects: (i) the precision used to analyse data and to detect the errors, before the data are published; (ii) the strategies of communication of results, which needs to be customized according to the audience to whom they are directed. It would be beneficial to create different tools for the stakeholders. In fact, policy makers may be interested in gaining a global overview of the territorial performance, together with a support for the decisions on the reorganization aspects; physicians may benefit from understanding their improvement through time, or the cause for worse results; patients may be willing to read in a clear and immediate way the characteristics of the hospitals to be chosen.

- **General Practitioners**. The role of the GP as a gate keeper is fostered by the integrated approach. In fact, if the integrated approach encourages patients to make their hospital choice, the GP is called to provide to his/her patients the tools to choose. The GP has to guide patients through the use of the information on quality performance that is published, in order to clarify the decisional elements that need to be evaluated by the patients themselves.

- **Patients**. The whole integrated approach, together with the importance given to the volume–outcome association, has the ultimate objective of improving patients' health conditions. Hence, patients represent the category that receives the highest gain from this new approach. More specifically, patients obtain the strengthening of two aspects: quality and freedom of choice. As for quality, they can rely on healthcare services with lower mortality. As for freedom of choice, they are encouraged to choose the hospital they prefer.

  However, patients need to be aware of this added value, and they also need to be able to take advantage from it. The territorial reconfiguration, from the patient perspective, causes indeed longer distances to be travelled and could cause longer waiting lists. Even though the two increases do not risk to endanger the health conditions, patients need to understand the reason behind these changes. The communication of the EBM strategies, together with the assistance of the GP, could help patients to comprehend that the unique priority is their health.

Chapter 6

Conclusions

Planning problems in healthcare systems have received great attention in the last decade. In fact, policy makers are required to provide a territorial reorganization of healthcare services, but, at the same time, to guarantee higher quality of clinical outcomes. Researchers in the health management field have investigated different strategies to improve the location and allocation of structural resources. However, they have rarely merged the managerial perspective with the clinical one, so that they have often focused primarily on organizational efficiency rather than on medical outcomes.

In this thesis, we studied the reorganization of hospitals operating in a territory, in particular, how to distribute volumes of activity among hospitals operating in a same geographical area. Our main contribution is twofold: (i) we have enriched the managerial approach with clinical insights; (ii) we have integrated the perspectives of two relevant stakeholders of the healthcare system that are usually considered individually, namely, the policy maker and the patient.

The merger between the managerial and the clinical points of view has been possible thanks to the use of the volume-outcome association, a recurring trend that associates the volume of activity (i.e., the number of interventions), which is a managerial parameter, with the outcome (e.g., mortality), which is a clinical parameter. Even though this association has been deeply documented in the medical literature, to the best of our knowledge this is the first time that it is practically applied to a planning problem.

The thesis is composed by two parts dedicated to the mentioned single stakeholders (i.e., policy maker and patient), and one related to their integration.

In the first part, by using the policy maker perspective, we developed a model to solve the planning problem on the allocation of interventions among hospitals, subject to epidemiological and structural constraints. Differently from previous researches that focused on economic or organizational efficiency, we have set as objective the minimization of the total mortality, thus focusing on a clinical measure of quality. As a consequence, our model provides a comparison among territorial configurations that is based on patients' health conditions. The use of quality as objective perfectly represents the duty of the policy maker, who is in charge of guaranteeing better health conditions to the whole population.

However, the policy maker model has been built on the assumption that patients will adhere to the strategical plan, i.e., patients are assumed to confirm volumes that have been strategically planned by the policy makers. The need for the prediction about patients' behaviours, who actually can cause the volume distribution to differ from the one planned by the policy maker, has led us to the second part of the thesis, focused on patients.

By using the patient perspective, we analysed patients' choice problem, i.e., the decision of each patient on which hospital to be treated in. Through the use of the well-know methodology of the conditional logit, we have represented the factors that affect the decisional process (both hospital and individual characteristics), and their impact on the utility gained by patients from choosing an alternative rather than another one. We have then applied the model to data on colon cancer patients from Piedmont that have been treated between 2004 and 2014. Few studies on choice models have focused so far on Italian patients, which raise interest since they have always been free to choose any hospital in the regional territory. Our results provide useful indications to local providers, which can be aware of the preferences of their patients, and of how these preferences change depending on patients' personal characteristics.

The development of the model of patient choice has been motivated not only by the interest in finding out how patients actually choose, but also by the need to integrate patients' behaviour in the policy maker's planning problem. In fact, even though the patient model is useful to realistically represent the freedom of choice that patients actually have, if used alone it omits the territorial overview of the service provided.

Hence, after having enlightened strengths and weaknesses of the individual decisional processes, we developed two approaches that integrate the two perspectives. In the first approach (the *quasi-integrated* approach), we have added to the policy maker perspective the consideration on the maximum distance that patients are willing to travel. By doing so, we have combined two areas of health studies that are usually considered as separated, i.e., health management and health economics. The result of the *quasi-integrated* approach is a territorial configuration that accounts for patient preferences. It is a very simple way of integrating patients and policy maker's perspectives; however, it is based on the assumption that patients will go to the hospital that has been thought for them, and no space is left to patients' own decisional process.

The second approach (the *fully-integrated* approach), instead, explicitly include the prediction of patient behaviour in the decisional process of the policy maker. It is a more complex approach, with respect to the *quasi-integrated* approach, but it is able to support planning decisions that (i) are effective in terms of better health outcomes, since the objective is still a clinical measure of quality through the use of the volume-outcome association; (ii) guarantee that patients' choices will respect the volumes that have been strategically planned, i.e., it integrates the perspectives of the two stakeholders.

Eventually, we have created a framework of key performance indicators, in order to explore the differences among the solutions reached by using all the proposed approaches, in terms of three quality measures, namely mortality, volume and distance. This framework represents a useful tool for the shareholders of the healthcare system, as it enables them to replicate the planning approaches to other territorial contexts without losing the ability to evaluate the results.

All the approaches have been applied to the same case study, and the framework of key performance indicators has been used to confirm the advantage of integrating the single actors' perspectives.

Policy implications of the integrated perspective have also been analysed, by deepening the consequences that the integrated approach has on all the stakeholders of the healthcare system, including providers, physicians, researchers, general practitioners and patients. The most important implication for the policy maker, who is the key actor of the planning problem, is that even though the new decisional approach requires more attention (i) to the production of scientific evidence that supports planning choices and (ii) to the incentives given in order to foster the change, it also acts as a guarantee of the increase in the population health conditions.

Several limitations apply to this research, related to additional internal and external factors that need to be considered. Internal aspects are those related to the surgical interventions: (i) the complexity of the treated cases, which make more appropriate to consider the volume–outcome association as stochastic rather than deterministic; (ii) the variability of the demand, which could change from one year to the following one, both in terms of total regional demand, and, most importantly, in terms of volumes allocated to hospitals; (iii) the identification of the over-treatment, which causes both the number of interventions to be higher and to alter the performance of the hospitals since they are dealing with not appropriate treatments.

External aspects, instead, concern accessibility to health care services, geographical characteristics of the catchment areas, costs incurred by providers [1] and the consideration of the perspectives of the other shareholders of the healthcare system (e.g., providers and physicians). Furthermore, in order to implement the strategical planning decision, the solution approaches should be developed over multiple years, with constantly updated information on hospital performance and scientific evidence [11]. Future research will be devoted to include these aspects.

References

[1] Laura Amato, Paola Colais, Marina Davoli, Eliana Ferroni, D Fusco, S Minozzi, F Moirano, P Sciattella, S Vecchi, M Ventura, et al. Volume and health outcomes: evidence from systematic reviews and from evaluation of italian hospital data. *Epidemiologia e prevenzione*, 37(2-3 Suppl 2):1–100, 2013.

[2] Agenas (National Agency for regional health services). *National Outcome Program.*, 2014. Available at http://95.110.213.190/PNEed14.

[3] Agenas (National Agency for regional health services). *National Outcome Program.*, 2013. Available at http://95.110.213.190/PNEed13.

[4] Harold S Luft, John P Bunker, and Alain C Enthoven. Should operations be regionalized? the empirical relation between surgical volume and mortality. *New England Journal of Medicine*, 301(25):1364–1369, 1979.

[5] David Archampong, David Borowski, Peer Wille-Jørgensen, and Lene H Iversen. Workload and surgeon s specialty for outcome after colorectal cancer surgery. *The Cochrane Library*, 2012.

[6] Nils Gutacker, Karen Bloor, Richard Cookson, Chris P Gale, Alan Maynard, Domenico Pagano, José Pomar, Enrique Bernal-Delgado, and as part of the ECHO collaboration. Hospital surgical volumes and mortality after coronary artery bypass grafting: using international comparisons to determine a safe threshold. *Health services research*, 52(2):863–878, 2017.

[7] John D Birkmeyer, Andrea E Siewers, Emily VA Finlayson, Therese A Stukel, F Lee Lucas, Ida Batista, H Gilbert Welch, and David E Wennberg. Hospital volume and surgical mortality in the united states. *New England Journal of Medicine*, 346(15):1128–1137, 2002.

[8] Ethan A Halm, Clara Lee, and Mark R Chassin. Is volume related to outcome in health care? a systematic review and methodologic critique of the literature. *Annals of internal medicine*, 137(6):511–520, 2002.

[9] Sudip K Sarker and Charles Vincent. Errors in surgery. *International Journal of Surgery*, 3(1):75–81, 2005.

[10] Agenas (National Agency for regional health services). *National Outcome Program.*, 2015. Available at http://95.110.213.190/PNEed15.

[11] Laura Amato, Danilo Fusco, Anna Acampora, Katia Bontempi, Alessandro Rosa, Paola Colais, Fabio Cruciani, et al. Volume and health outcomes: evidence from systematic reviews and from evaluation of italian hospital data. *Epidemiologia e prevenzione*, 41(5-6 Suppl 2):1–128, 2017.

[12] Teodor P Grantcharov, Linda Bardram, Peter Funch-Jensen, and Jacob Rosenberg. Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills. *The American journal of surgery*, 185(2):146–149, 2003.

[13] Italian Health Ministry. *National Health Service, principles.*, 2018. available on-line http://www.salute.gov.it/.

[14] Tillal Eldabi, Zahir Irani, and Ray J Paul. A proposed approach for modelling health-care systems for understanding. *Journal of Management in Medicine*, 16(2/3):170–187, 2002.

[15] Christine Pelletier and Georges Weil. Defining the regional healthcare planning objective using a multi-criteria approach. *Logistics Information Management*, 16(3):191–200, 2003-06.

[16] Abdur Rais and Ana Viana. Operations research in healthcare: a survey. *International transactions in operational research*, 18(1):1–31, 2011.

[17] Mark S. Daskin and Latoya K. Dean. Location of health care facilities. In Margaret L. Brandeau, François Sainfort, and William P. Pierskalla, editors, *Operations Research and Health Care*, volume 70, pages 43–76. Kluwer Academic Publishers, 2005.

[18] Gustavo Daniel Buzai. *Location-allocation models applied to urban public services. Spatial analysis of Primary Health Care Centers in the city of Luján, Argentina.* Hungarian Academy of Sciences. Research Centre for Astronomy and Earth Sciences. Geographical Institute, 2013.

[19] Tom Storme and Frank Witlox. Location-allocation models. In Douglas Richardson, Noel Castree, Michael F. Goodchild, Audrey Kobayashi, Weidong Liu, and Richard A. Marston, editors, *International Encyclopedia of Geography: People, the Earth, Environment and Technology*, pages 1–4. John Wiley & Sons, Ltd, 2017-03-06.

[20] JC Ridge, SK Jones, MS Nielsen, and AK Shahani. Capacity planning for intensive care units. *European journal of operational research*, 105(2):346–355, 1998.

[21] P. R. Harper and A. K. Shahani. Modelling for the planning and management of bed capacities in hospitals. *The Journal of the Operational Research Society*, 53(1):11–18, 2002.

[22] Nelly Litvak, Marleen van Rijsbergen, Richard J. Boucherie, and Mark van Houdenhoven. Managing the overflow of intensive care patients. *European Journal of Operational Research*, 185(3):998–1010, 2008-03.

[23] Ali Kokangul. A combination of deterministic and stochastic approaches to optimize bed capacity in a hospital unit. *Computer Methods and Programs in Biomedicine*, 90(1):56–65, 2008-04.

[24] Yue Zhang, Oded Berman, and Vedat Verter. Incorporating congestion in preventive healthcare facility network design. *European Journal of Operational Research*, 198(3):922–935, 2009-11.

[25] Jeffery K. Cochran and Aseem Bharti. Stochastic bed balancing of an obstetrics hospital. *Health Care Management Science*, 9(1):31–45, 2006-02.

[26] Linda V. Green and Sergei Savin. Reducing delays for medical appointments: A queueing approach. *Operations Research*, 56(6), 2008.

[27] Maartje E. Zonderland, Richard J. Boucherie, Nelly Litvak, and Carmen L. A. M. Vleggeert-Lankamp. Planning and scheduling of semi-urgent surgeries. *Health Care Management Science*, 13(3):256–267, 2010-09.

[28] L Mayhew and D Smith. Using queuing theory to analyse the government's 4-h completion time target in accident and emergency departments. *Health care management science*, 11(1):11–21, 2008.

[29] Ivo Adan, Jos Bekkers, Nico Dellaert, Jan Vissers, and Xiaoting Yu. Patient mix optimisation and stochastic resource requirements: A case study in cardiothoracic surgery planning. *Health Care Management Science*, 12(2):129–141, 2009-06.

[30] Peter T. VanBerkel and John T. Blake. A comprehensive simulation for wait time reduction and capacity planning applied in general surgery. *Health Care Management Science*, 10(4):373–385, 2007-10-16.

[31] William S. Lovejoy and Ying Li. Hospital operating room capacity expansion. *Management Science*, 48(11):1369–1387, 2002.

[32] Jeroen Beliën and Erik Demeulemeester. Building cyclic master surgery schedules with leveled resulting bed occupancy. *European Journal of Operational Research*, 176(2):1185–1204, 2007-01.

[33] Alessandro Agnetis, Alberto Coppi, Matteo Corsini, Gabriella Dellino, Carlo Meloni, and Marco Pranzo. Long term evaluation of operating theater planning policies. *Operations Research for Health Care*, 1(4):95–104, 2012.

[34] Jeroen M van Oostrum, Mark Van Houdenhoven, Johann L Hurink, Erwin W Hans, Gerhard Wullink, and Geert Kazemier. A master surgical scheduling approach for cyclic scheduling in operating room departments. *OR spectrum*, 30(2):355–374, 2008.

[35] Ruth E. Wachtel and Franklin Dexter. Tactical increases in operating room block time for capacity planning should not be based on utilization:. *Anesthesia & Analgesia*, 106(1):215–226, 2008-01.

[36] Guoxuan Ma and Erik Demeulemeester. A multilevel integrative approach to hospital case mix and capacity planning. *Computers & Operations Research*, 40(9):2198–2207, 2013-09.

[37] Vedat Verter and Sophie D. Lapierre. Location of preventive health care facilities. page 10, 2002.

[38] Qingjin Peng Hamid Afshari. Challenges and solutions for location of healthcare facilities. *Industrial Engineering & Management*, 03(2), 2014.

[39] Shams-ur Rahman and David K. Smith. Use of location-allocation models in health service development planning in developing nations. *European Journal of Operational Research*, 123(3):437–452, 2000-06.

[40] Stephen Mahar, Kurt M Bretthauer, and Peter A Salzarulo. Locating specialized service capacity in a multi-hospital network. *European Journal of Operational Research*, 212(3):596–605, 2011.

[41] Khadijeh Naboureh and Ehram Safari. A stochastic location-allocation model for specialized services in a multihospital system. *Advances in Operations Research*, 2016:1–16, 2016.

[42] Wei Gu, Xin Wang, and S. Elizabeth McGregor. Optimization of preventive health care facility locations. *International Journal of Health Geographics*, 9(1):17, 2010.

[43] Dong-Guen Kim and Yeong-Dae Kim. A branch and bound algorithm for determining locations of long-term care facilities. *European Journal of Operational Research*, 206(1):168–177, 2010-10.

[44] Hongzhong Jia, Fernando Ordóñez, and Maged Dessouky. A modeling framework for facility location of medical services for large-scale emergencies. *IIE Transactions*, 39(1):41–55, 2007-01.

[45] Roberto D. Galvão, Luis Gonzalo Acosta Espejo, Brian Boffey, and Derek Yates. Load balancing and capacity constraints in a hierarchical location model. *European Journal of Operational Research*, 172(2):631–646, 2006-07.

[46] Maria Elena Bruni, Domenico Conforti, Nicola Sicilia, and Sandro Trotta. A new organ transplantation location–allocation policy: a case study of italy. *Health Care Management Science*, 9(2):125–142, 2006-05.

[47] Eyüp Çetin. A blood bank location model: A multiobjective approach. *Eur. J. Pure Appl. Math*, page 13, 2009.

[48] S.S. Radiah Shariff, Noor Hasnah Moin, and Mohd Omar. Location allocation modeling for healthcare facility planning in malaysia. *Computers & Industrial Engineering*, 62(4):1000–1010, 2012-05.

[49] Ping Yin and Lan Mu. Modular capacitated maximal covering location problem for the optimal siting of emergency vehicles. *Applied Geography*, 34:247–254, 2012-05.

[50] Panagiotis Mitropoulos, Ioannis Mitropoulos, Ioannis Giannikos, and Aris Sissouras. A biobjective model for the locational planning of hospitals and health centers. *Health Care Management Science*, 9(2):171–179, 2006-05.

[51] Jomon Aliyas Paul. Models for hospital location and capacity allocation for an area prone to natural disasters. page 24, 2008.

[52] Christian Stummer, Karl Doerner, Axel Focke, and Kurt Heidenberger. Determining location and size of medical departments in a hospital network: A multiobjective decision support approach. *Health Care Management Science*, 7(1):63–71, 2004-02.

[53] P Harper. Planning health services with explicit geographical considerations: a stochastic location?allocation approach. *Omega*, 33(2):141–152, 2005-04.

[54] Karsten Vrangbaek, Ruth Robertson, Ulrika Winblad, Hester Van de Bovenkamp, and Anna Dixon. Choice policies in northern european health systems. *Health Economics, Policy and Law*, 7(1):47–71, 2012-01.

[55] Timothy Besley and Maitreesh Ghatak. Incentives, choice, and accountability in the provision of public services. *Oxford Review of Economic Policy*, 19(2):235–249, 2003.

[56] Giuseppe Moscelli, Luigi Siciliani, Nils Gutacker, and Hugh Gravelle. Location, quality and choice of hospital: Evidence from england 2002–2013. *Regional Science and Urban Economics*, 60:112–124, 2016-09.

[57] Walter Beckert, Mette Christensen, and Kate Collyer. Choice of nhs-funded hospital services in england. *The Economic Journal*, 122(560):400–417, 2012-05.

[58] Nils Gutacker, Luigi Siciliani, Giuseppe Moscelli, Hugh Gravelle, et al. Do patients choose hospitals that improve their health? Technical report, Centre for Health Economics, University of York, 2015.

[59] Rita Santos, Hugh Gravelle, and Carol Propper. Does quality affect patients' choice of doctor? evidence from england. *The Economic Journal*, 127(600):445–494, 2017-03.

[60] Abigail Tay. Assessing competition in hospital care markets: the importance of accounting for quality differentiation. *RAND Journal of Economics*, pages 786–814, 2003.

[61] Harold Luft, S., Deborah Garnick, David Mark, Deborah Peltzman, Claran Phibbs, Erik Lichtenberg, and Stephen Mcphee. Does quality influence choice of hospital? *JAMA: The Journal of the American Medical Association*, 263(21), 1990-06-06.

[62] Janet M. Bronstein and Michael A. Morrisey. Bypassing rural hospitals for obstetrics care. *Journal of Health Politics, Policy and Law*, 16(1):87–118, 1991.

[63] Katherine Ho. The welfare effects of restricted hospital choice in the US medical care market. *Journal of Applied Econometrics*, 21(7):1039–1079, 2006-11.

[64] F. Moscone, E. Tosetti, and G. Vittadini. Social interaction in patients' hospital choice: evidence from italy: Social interaction in patients' hospital choice. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(2):453–472, 2012-04.

[65] Marco Varkevisser and Stéphanie A. van der Geest. Why do patients bypass the nearest hospital? an empirical analysis for orthopaedic care and neurosurgery in the netherlands. *The European Journal of Health Economics*, 8(3):287–295, 2007-09.

[66] Marco Varkevisser, Stéphanie A. van der Geest, and Frederik T. Schut. Do patients choose hospitals with high quality ratings? empirical evidence from the market for angioplasty in the netherlands. *Journal of Health Economics*, 31(2):371–378, 2012-03.

[67] Puck D. C. Beukers, Ron G. M. Kemp, and Marco Varkevisser. Patient hospital choice for hip replacement: empirical evidence from the netherlands. *The European Journal of Health Economics*, 15(9):927–936, 2014-12.

[68] Noritoshi Yoshii and Kiyohide Fushimi. *Factors inuencing the selection of a hospital for colon cancer surgery in Japan: analysis of the effects of surgery volume, hospital functions, and geographic conditions.*, 2006.

[69] Aafke Victoor, Diana MJ Delnoij, Roland D Friele, and Jany JDJM Rademakers. Determinants of patient choice of healthcare providers: a scoping review. *BMC Health Services Research*, 12(1), 2012-12.

[70] Sherman T. Folland. Predicting hospital market shares. *Inquiry*, 20(1):34–44, 1983.

[71] Gauthier Bouche, Virginie Migeot, Simone Mathoulin-Pélissier, Roger Salamon, and Pierre Ingrand. Breast cancer surgery: Do all patients want to go to high-volume hospitals? *Surgery*, 143(6):699–705, 2008-06.

[72] Lawton R. Burns and Douglas R. Wholey. The impact of physician characteristics in conditional choice models for hospital care. *Journal of Health Economics*, 11(1):43–62, 1992-05.

[73] Martin Gaynor, Carol Propper, and Stephan Seiler. Free to choose? reform, choice, and consideration sets in the english national health service. *American Economic Review*, 106(11):3521–3557, 2016-11.

[74] Peter Sivey. The effect of waiting time and distance on hospital choice for english cataract patients. *Health Economics*, 21(4):444–456, 2012-04.

[75] Martin Gaynor, Carol Propper, and Stephan Seiler. Free to choose? reform and demand response in the english national health service. Technical report, National Bureau of Economic Research, 2012.

[76] Suzanne Ruwaard and Rudy Douven. Quality and hospital choice for cataract treatments: the winner takes most. page 46, 2014.

[77] Devin G. Pope. Reacting to rankings: Evidence from “america’s best hospitals”. *Journal of Health Economics*, 28(6):1154–1165, 2009-12.

[78] David H Howard. Quality and consumer choice in healthcare: Evidence from kidney transplantation. page 22, 2005.

[79] Dominic Hodgkin. Specialized service offerings and patients’ choice of hospital: The case of cardiac catheterization. *Journal of Health Economics*, 15(3):305–332, 1996-06.

[80] MM Chowdhury, H Dagash, and A Pierro. A systematic review of the impact of volume of surgery and specialization on patient outcome. *British Journal of Surgery: Incorporating European Journal of Surgery and Swiss Surgery*, 94(2):145–161, 2007.

[81] Johannes Morche, Tim Mathes, and Dawid Pieper. Relationship between surgeon volume and outcomes: a systematic review of systematic reviews. *Systematic reviews*, 5(1):204, 2016.

[82] Russell L Gruen, Veronica Pitt, Sally Green, Anne Parkhill, Donald Campbell, and Damien Jolley. The effect of provider case volume on cancer mortality: systematic review and meta-analysis. *CA: a cancer journal for clinicians*, 59(3):192–211, 2009.

[83] Ingemar Ihse. The volume-outcome relationship in cancer surgery: a hard sell. *Annals of surgery*, 238(6):777, 2003.

[84] Teodor P Grantcharov, Linda Bardram, Peter Funch-Jensen, and Jacob Rosenberg. Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills. *The American journal of surgery*, 185(2):146–149, 2003.

[85] Harold S Luft, Sandra S Hunt, and Susan C Maerki. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? *Health services research*, 22(2):157, 1987.

[86] Martin Gaynor, Harald Seider, and William B Vogt. The volume-outcome effect, scale economies, and learning-by-doing. *American Economic Review*, 95(2):243–247, 2005.

[87] Roos Mesman, Gert P Westert, Bart JMM Berden, and Marjan J Faber. Why do high-volume hospitals achieve better outcomes? a systematic review about intermediate factors in volume–outcome relationships. *Health Policy*, 119(8):1055–1067, 2015.

[88] John D Birkmeyer and Justin B Dimick. Potential benefits of the new leapfrog standards: effect of process and outcomes measures. *Surgery*, 135(6):569–575, 2004.

[89] Arnold M Epstein. Volume and outcome—it is time to move ahead, 2002.

[90] Arianna Alfieri, Elisabetta Listorti, and Andrea Matta. Improving surgical outcomes through optimal volumes allocation. *Flexible Services and Manufacturing Journal*, pages 1–24, 2016.

[91] Jan Mainz. Defining and classifying clinical indicators for quality improvement. *International Journal for Quality in Health Care*, 15(6):523–530, 2003.

[92] Paula Braveman and Sofia Gruskin. Defining equity in health. *Journal of Epidemiology & Community Health*, 57(4):254–258, 2003.

[93] *Choosing Wisely.*, 2008. Available on-line at http://www.choosingwisely.org/.

[94] Philip E Gill and Elizabeth Wong. Sequential quadratic programming methods. In *Mixed integer nonlinear programming*, pages 147–224. Springer, 2012.

[95] Department of Epidemiology ASLTO3 Piedmont Region. *MADE Demographic and Epidemiologic Search Engine database on the Internet.*, 2015. Available at http://www.ruparpiemonte.it/.

[96] Rosella Levaggi and Marcello Montefiori. *Health care provision and patient mobility*. Springer, 2014.

[97] *Decreto Ministeriale 2 aprile 2015 n.70, Regolamento recante definizione degli standard qualitativi, strutturali, tecnologici equantitativi relativi all'assistenza ospedaliera.*, 2015. available on-line http://www.salute.gov.it/.

[98] *Rapporto sull'attivita' del ricovero ospedaliero, Dati SDO primo semestre 2014.*, 2014. available on-line http://www.salute.gov.it/.

[99] George France, Francesco Taroni, and Andrea Donatini. The italian health-care system. *Health economics*, 14(S1):S187–S202, 2005.

[100] Daniel McFadden et al. Conditional logit of qualitative choice behavior. 1974.

[101] Karin Monstad, Lars Birger Engesæter, and Birgitte Espehaug. Patients' preferences for choice of hospital. *Health Economics Bergen Working Paper*, 5(06), 2006.

[102] Saul D Hoffman and Greg J Duncan. Multinomial and conditional logit discrete-choice models in demography. 25(3):415–427, 1988.

[103] AIRTUM AIOM. I numeri del cancro in italia. *Roma: Il Pensiero Scientifico Editore*, 2016.

[104] Giacomo Ruffo, Giuliano Barugola, Roberto Rossini, and Carlo Augusto Sartori. Colorectal surgery in italy. criteria to identify the hospital units and the tertiary referral centers entitled to perform it. *Updates in surgery*, 68(2):123–128, 2016.

[105] Agenas (National Agency for regional health services). *National Outcome Program.*, 2016. Available at http://95.110.213.190/PNEed16.

[106] Piedmont Region. *Deliberazione della Giunta Regionale 23 novembre 2015*, 2015.

[107] Piedmont Region. *Screening oncologici: Prevenzione Serena*, 2017. Available at http://www.regione.piemonte.it/sanita/cms2/guida-al-servizio-sanitario/prevenzione/275-screening-oncologici.

[108] Anne Elixhauser, Claudia Steiner, D Robert Harris, and Rosanna M Coffey. Comorbidity measures for use with administrative data. *Medical care*, pages 8–27, 1998.

[109] IRES Piemonte (Sylvie Occelli). *Matrice delle distanze fra i Comuni Piemontesi*. Available in Banche Dati della rete dei servizi di epidemiologia del Piemonte.

[110] Giuliana De Luca, Alessio Petrelli, Tania Landriscina, Roberto Gnavi, Massimo Giammaria, and Giuseppe Costa. Geographic and socioeconomic differences in access to revascularization following acute myocardial infarction. *The European Journal of Public Health*, 26(5):760–765, 2016.